DOCUMENT RESUME

ED 068 455 24 SP 005 914

AUTHOR Olson, David R.; And Others
TITLE Teacher-Student Communication Games: Some Experiments on Instruction.
INSTITUTION Stanford Univ., Calif. Stanford Center for Research and Development in Teaching.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
BUREAU NO BR-5-0252
PUB DATE Sep 72
CONTRACT OEC-6-10-078
NOTE 102p.; Research and Development Memorandum No. 94

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS Classroom Communication; Discovery Learning; *Educational Games; *Instruction; *Learning Processes; Learning Theories; Research; *Student Teacher Relationship; *Verbal Communication

ABSTRACT

This inquiry began with the observation that learning from instruction is radically more efficient for obtaining information than learning by discovery. A series of seven experiments was conducted to determine some of the factors involved in learning from verbal instruction. The perspective adopted was that of communication theory, in which the teacher is considered an encoder who tries to generate a message that will produce a desired effect on the child, who acts as the decoder of the message. The experiments studied were directed to such questions as the following: Do teachers differ in a consistent way in their skill in generating instructional statements? What are the dimensions on which they differ? Do these differences have a consistent effect on the student outcomes? What is the role of feedback in this process? The experimental results were considered in terms of three criteria: the substantive findings, the method of research or paradigm developed, and the practical application of this research to improving instruction. An extensive bibliography is included. (Author/MJM)

ED 068455

STANFORD CENTER
FOR RESEARCH AND DEVELOPMENT
IN TEACHING

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Research and Development Memorandum No. 94

TEACHER-STUDENT COMMUNICATION GAMES: SOME EXPERIMENTS ON INSTRUCTION

David R. Olson, Donna Crossan Case,
William Wine, in collaboration with
Susan Pagliuso, Diana Crosbie, A. J. H.
Gaite

School of Education
Stanford University
Stanford, California

September 1972

SP 005 914

Published by the Stanford Center for Research and Development in Teaching, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare (Contract No. OEC-6-10-078). The opinions expressed in this publication do not necessarily reflect the position or policy of the Office of Education, and no official endorsement by the Office of Education should be inferred.

Introductory Statement

The Center's mission is to improve teaching in American schools. Too many teachers still employ a didactic style aimed at filling passive students with facts. The teacher's environment often prevents him from changing his style, and may indeed drive him out of the profession. And the children of the poor typically suffer from the worst teaching.

The Center uses the resources of the behavioral sciences in pursuing its objectives. Drawing primarily upon psychology and sociology, but also upon other behavioral science disciplines, the Center has formulated programs of research, development, demonstration, and dissemination in three areas. Program 1, Teaching Effectiveness, is now developing a Model Teacher Training System that can be used to train both beginning and experienced teachers in effective teaching skills. Program 2, The Environment for Teaching, is developing models of school organization and ways of evaluating teachers that will encourage teachers to become more professional and more committed. Program 3, Teaching Students from Low-Income Areas, is developing materials and procedures for motivating both students and teachers in low-income schools.

The present report deals with work performed at the Ontario Institute for Studies in Education, and at the Center, by David Olson and his co-workers. Dr. Olson wrote much of this work while he was a Visiting Scholar at the Center during the academic year 1969-1970. This work bears most directly on the Explaining Project of the Technical Skills Activity within the Training Systems Design Component of the Program on Teaching Effectiveness.

Acknowledgments

The research reported here was supported by a grant (No. 0231) from the Research Review Board of the Ontario Institute for Studies in Education, Toronto, Ontario. The study was carried out with the cooperation of the York Borough Board of Education, including the faculties and students of King George Junior Public School, Weston Memorial Public School, Humewood Public School, H. J. Alexander Public School, and York Memorial High School.

The project benefited from the advice of Kent Henderson and Ross Traub and the assistance of Carolyn Dundas. It was indirectly supported by a Canada Council Fellowship and a study leave granted to the senior author by OISE. The generosity of the Stanford Center for Research and Development in Teaching in providing the senior author with an intellectual home, editorial and publishing services, and the secretarial services of Unarose Thompson, Lupria Baker, and Ann Gladstone, is also gratefully acknowledged.

D. R. O.

About the Authors

David R. Olson is Professor of Applied Psychology at the Ontario Institute for Studies in Education in Toronto, Canada. Donna Crossan Case was a Research Assistant at the Institute. William Wine, formerly a student at the Institute, is now at the Clark Institute of Psychiatry, Toronto. Susan Pagliuso, a Research Assistant on this project, is now in the Department of Curriculum at the Institute. Diana Crosbie was a student at the Institute. A. J. H. Gaite is Assistant Professor of Educational Psychology at the University of Wisconsin.

Contents

List of Tables vii

List of Figures viii

Abstract ix

Introduction 1

Experiment 1: Encoder Variability 6

- Subjects 6
- Materials 6
- Procedures 6
- Results and Discussion 7

Experiment 2: The Development of Optimal Messages 10

- Subjects 11
- Messages 11
- Procedures 12
- Results and Discussion 12

Experiment 3: Some Other Aspects of Optimal Messages 16

- Subjects 16
- Materials 16
- Procedures 16
- Results and Discussion 16

Experiment 4: The Effects of Recoding and Pupil Ability on Decodability 19

- Subjects 21
- Materials 21
- Procedures 25
- Results 25
- Discussion 32

Experiment 5: Age of Decoder as a Variable in Determining the Effectiveness of a Verbal Message 37

- Subjects 38
- Experimental Lists 38
- Procedures 39
- Results 39
- Discussion 46

Experiment 6: The Teacher's Ability to Encode Messages 48

Subjects 48
Materials 48
Procedures 50
Results and Discussion 51
Conclusions 65

Experiment 7: The Role of Feedback in Teacher-Child Communication 71

Subjects 74
Experimental Groups 74
Materials 75
Procedures 75
Results and Discussion 76
Conclusions 81

General Summary, Conclusions, and Implications 83

References 89

List of Tables

1-1. Message Length and Decodability 7

2-1. Analysis of Variance for the Dimensions Correctly Decoded for the Basic and Elaborated Messages 14

3-1. Analysis of Variance for Decoding by Twenty Children of Messages with Covered and Uncovered Forms 17

4-1. One-Way Analysis of Variance of the Number of Words Recalled for the Seven Treatment Groups 27

4-2. Comparison of Word-List Recall after Seven Forms of Instruction (Scheffé Mothod) 27

4-3. One-Way Analysis of Variance of Memory-for-Discrete-Instructions Scores under Three Instructional Conditions . 30

4-4. Comparison of Recall after Three Forms of Instruction on Memory for Discrete Instructions (Scheffé Method) . . . 31

6-1. Four-Way Analysis of Variance for the Effects of Encoder, Decoder, Grade Level of Decoder, and Order of Tasks on the Decodability of Verbal Messages 52

6-2. Four-Way Analysis of Variance for the Effects of Encoder, Decoder, Grade Level of Decoder, and Order of Tasks on the Decodability of Geometric Messages 54

6-3. Correlation Matrix for Sixteen Encoder, Message, and Decoder Variables 57

6-4. Correlation Matrix for IQ and Individual Decoder Score on Verbal and Geometric Tasks 58

7-1. Two-Way Analysis of Variance of Auditory and Visual Feedback 78

7-2. Multiple Correlation Coefficients for Encoder IQ, Decoder IQ, and Communication Score under Four Feedback Conditions 79

List of Figures

1. Smith's model of teaching 4

2-1. Percentage of binary features correct over four problems for basic and elaborated messages 13

4-1. Mean number of words recalled as a function of encoding condition 26

4-2. Number of words recalled as a function of digit-span score: Regression analyses 28

4-3. Mean score on the memory-for-discrete-instruction test as a function of encoding condition 30

4-4. Regression lines for memory-for-discrete-instructions score as a function of digit-span score 32

5-1. The effects of ordering on recall as a function of sex and grade level 40

5-2. The effects of ordering on recall as a function of number of trials 41

5-3. The effects of superordinate terms on recall as a function of sex, grade level, and number of trials 42

5-4. Mean number of words recalled on Trial 4 as a function of sex, grade, ordering, and superordination 43

5-5. Percentage of superordinate terms generated by S as a function of treatment condition, sex, and grade level . . 44

6-1. Teacher's model for the geometric task 49

6-2. Student's answer sheet for the geometric task 49

6-3. Grade-order interaction on the verbal task 53

6-4. Grade-order interaction on the geometric task 55

7-1. Intended message, received message, and feedback 72

7-2. Communication effectiveness under four conditions of feedback as measured by decoder performance 77

8-1. A schematic model of the teacher-student communication system 84

Abstract

This inquiry began with the observation that learning from instruction is a radically more efficient means of obtaining information than learning by discovery. On the assumption that these two processes differ not only in their effectiveness but in their functioning, a series of seven experiments was conducted to determine some of the factors involved in learning from verbal instruction. The perspective adopted was that of communication theory, in which the teacher is considered an encoder who tries to generate a message that will produce a desired effect on the child, who acts as the decoder of the message.

The experiments studied were directed to such questions as the following: Do teachers differ in a consistent way in their skill in generating instructional statements? What are the dimensions on which they differ? Do these differences have a consistent effect on the student outcomes? What is the role of feedback in this process?

The experimental results were considered in terms of three criteria: the substantive findings themselves, the method of research or paradigm developed here, and the practical application of this research to improving instruction.

INTRODUCTION

Linus Pauling worked on the problem of the molecular structure of protein intermittently for over a decade--from 1937 until he finally discovered the alpha helix and the gamma helix in 1948. He says:

> During the summer of 1937 I devoted a good part of my time, perhaps half-time for a few weeks, to the attempt to find a way of folding polypeptide chains that would account for the X-ray pattern of keratin. . . . I worked a little on the problem during succeeding years, and, of course, we had a large program of experimental investigation of simpler substances, amino acids and peptides, going on. I think that it took me about a day to discover the alpha helix and gamma helix in the spring of 1948, when I settled down to work on the problem again [personal communication, 1970].

Yet he was able to describe his discovery to the scientific community in a brief report that could be read in about 15 minutes. What Pauling had taken years to learn, in other words, his colleagues and students learned in minutes by the radically efficient means we call instruction. Thus learning by experience or discovery may be contrasted with learning by instruction--a distinction made centuries ago, for that matter, by Saint Thomas Aquinas: "There are two ways of acquiring knowledge. In one way, natural reason by itself reaches knowledge of unknown things, and this way is called discovery; in the other way, someone else aids the learner's natural reason, and this is called learning by instruction ["De Veritate," in Price, 1967, p. 197].

Realizing not only that instruction was important but that it relied heavily on language, both Saint Thomas and his intellectual forbear Saint Augustine proceeded to examine the nature of language. But finding that language consisted primarily of signs, the comprehension of which depended on a prior knowledge of reality, they concluded that little could be learned directly from the language of the teacher. Augustine, for example, stated that

> people deceive themselves in calling persons "teachers" who are not such at all, merely because generally there is no interval between the time of speaking and the time of knowing. And because they are quick to learn internally following the prompting of the one who speaks, they think they have learned externally from the one who was only a prompter ["De Magistro," in Price, 1967, p. 159].

Whatever the precise relationship between the teacher's language and the child's acquisition of knowledge, the schools have been charged with the task of arranging instructional means for transmitting to each succeeding generation of children the accumulated discoveries and conventions of the human species--that is, the culture. But how is this instruction accomplished? What conditions are necessary for instruction to occur? What makes one means of instruction more effective than another? How many dif-

ferent means are there to choose from? What awareness of these different means and what skill in choosing among them do teachers have?

Although considerable research has been devoted to how various species learn by experience (i.e., by discovery), less attention has been paid to how one learns from instruction and to the teacher's role in the instructional process. For example, the study of teacher effectiveness has centered not on what a teacher says or does--that is, on the information that is communicated, and how it is communicated most effectively--but on a set of personality characteristics presumed to have some bearing on the teacher's ultimate performance in the classroom. Travers (1963), after reviewing this research, has concluded that a half-century of it has not led to any clear formulation of what makes a teacher effective. Other reviewers have tended to agree (e.g., Ackerman, 1954).

Furthermore, the psychology of learning has never successfully explained in any systematic fashion the relationship between what one learns from direct experience and what one learns from instruction; nor has it explored the possibility that learning proceeds quite differently in these two cases. "Instruction" in experimental psychology is usually considered simply part of the preparation for the experiment proper. Yet according to Premack and Schwartz (1966), for example, "Organisms without language can be trained, but those with language can be instructed. We see instruction as presently resisting completely all attempts at stimulus response explication [p. 298]."

The variables that may be relevant to the effectiveness of instruction have not been simply ignored, of course. But in most accounts they have been deduced directly from learning theory, a questionable procedure if learning from instruction and learning from experience are in fact different processes. At any rate, the factors thought to be most critical in determining the effectiveness of instruction include, according to Bloom (1968), the clarity of the cues, practice, and reinforcements, and according to Ausubel (1963), the cognitive structure of the learner, the meaningfulness of the material, the frequency and distribution of practice, the learner's awareness of results (feedback), and the size, difficulty, and organization of the task. An extensive review of the factors that have been empirically implicated has recently been made by Gagné and Rohwer (1969) in terms of a set of conditions that may be manipulated by an instructor. Their list includes the gaining attention, establishing a "set," presenting the stimuli, using language for guiding and prompting, eliciting responses, and providing feedback to the learner. In their account, language is used primarily to establish a "set," as in giving instructions prior to an experiment, and to prompt and guide the learner. As a means of presenting stimuli, they consider words inferior to "pictorial materials [p. 394]." But as Olson (1970c) has pointed out, this apparent superiority of pictures may be due to the fact that pictures are more ambiguous than words, and whereas that ambiguity may be useful in the recall of verbal lists, it is a disadvantage in tasks requiring transfer.

However useful such conclusions may be, none seem to shed much light on the central problem attacked not only by Augustine but by contemporary

philosophers of language and some psychologists. This problem is, of course, the relationship between words and reality, or to put it another way, why language can be used to describe reality, as in instruction. Ausubel (1968) has at least made his assumption explicit; namely, the assumption that the word evokes the same cognitive content as the corresponding event (p. 42). Unfortunately, the assumption appears to be false in view of the evidence mentioned above (Gagné & Rohwer, 1969; Olson 1970c) which showed that words, pictures, and objects are not cognitively equivalent.

The structure of information in various media, such as language, pictures, and film, while it is also of central significance for a theory instruction, is likewise ignored by the present studies. Instead, researchers have so far examined the relative effectiveness of a variety of types of verbal messages as instructional means and teachers' skills in using them.

Another line of research relevant to the use of language in instruction comes from the development of communication theory (Shannon & Weaver, 1949), an attempt to provide a model for the flow of information from a source to a receiver. In a related development, Katz and Lazarsfeld (1955) summarized the flow of information from the mass media to the public in terms of four primary variables: the audience exposure, the differential character of the media, the content or subject matter of the media, and the characteristics of the audience. Furthermore, to account for the effects of the mass media, they postulated a "two-step flow of communication" in which information leaders essentially mediated between the information conveyed by the media and the opinions formed by the listeners. This conception of communication as the transmission from a source or encoder to a decoder has been generalized to interpersonal interactions by Berlo (1960) and more recently to teaching.

Although there are several such models of teaching (Gage, 1963), they all involve a common cycle including the teacher's perception of the situation, his message, the child's perception of the teacher and the message, and the child's reaction. The model developed by Smith (1960) is typical and can serve as a starting point for our conceptualization of the instructional process. Essentially it is a transmission model in which the teacher's performance of a teaching act causes some change in the performance of the child; seeing the child's act, the teacher then goes on to perform a new teaching act. Thus there is a cycle of teaching and learning, giving and taking instruction. This model in a simplified form is depicted in Figure I-1.

Smith's approach led him to describe classroom interaction in terms of a taxonomy of teachers' actions (Smith, Meux, Coombs, Nuthall, & Precians, 1967). Subsequently Nuthall (1968), using the different teaching strategies isolated by Smith, tried to generate ways of teaching concepts, such as giving instances of a concept rather than describing it. Some of Smith's strategies appeared to be more effective than others, at least for some concepts. Nuthall concluded that further progress in understanding

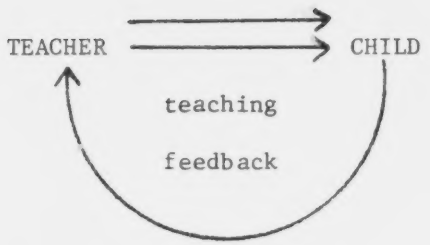

Fig. 1. Smith's model of teaching.

the effects of various strategies might require a more extensive analysis of the nature of the concepts to be taught, since some strategies might be more appropriate for some types of concepts than others. Smith's model, nonetheless, served as a starting point for describing the process of school instruction.

Many other studies have attempted to describe the process of classroom instruction (see Simon & Boyer, 1966-70). This approach has produced some interesting findings to be sure, but it appears to be based on the empiricist view that correct theories are discovered simply by looking intently at nature. Unfortunately, in order to make correct observations one must not only look intently but know what to look for. As Miller (1962) has pointed out, the probability of observing the right things by chance is negligible. If the categories for observation have little or no theoretical basis, they may well be superficial. On hearing the claim that the world consists of officers, serving maids, and chimney-sweeps, Kirkegaard responded that such a list had the good effect of setting the mind in motion but led one to overlook the fact that the categories were completely arbitrary (Murdoch, 1953). To avoid this pitfall, the studies reported here have relied for the most part on examining the effects on learners of various types of activities carried out by teachers. The system chosen for describing the teachers' acts is the one that most economically accounts for the differential learning of the children.

Conceptualizing the strategies used by the teacher therefore remains problematic. Joyce and Harootunian (1967) suggested that the strategy chosen by the teacher must be closely tied to the objectives of instruction: "A teacher who cannot readily vary his method or style is seriously limited. He needs to be able to select from a repetoire, tactics that will lead to different objectives and induce different students to learn [p. 95]." Which tactics or, as we prefer to say, forms of instruction, are most effective in reaching different types of objectives is one of the first concerns of this report. The use of different strategies for different learners is considered later.

Consider the direct access to the process and effects of instruction, particularly verbal instruction, that is offered by communication theory

(Berlo, 1960). In the simplest form of this theory, the teacher as encoder has the responsibility of formulating a message that can be assimilated or comprehended by the child, the decoder. Obviously the decoding may fail to occur, for any of several reasons. For example, the message may not contain all of the necessary information, or it may be organized in such a way that the child has difficulty extracting the information from it. Furthermore, to understand the message the child must know the meanings of words and phrases used in it, or in Wittgenstein's (1958) terms, the uses of words in various language games. Perhaps most important is the fact that a sentence is informative only to the extent that it partitions or specifies alternatives being considered by the listener. Thus a message that is ambiguous to one child may or may not be ambiguous to another child who is considering different alternatives (Olson, 1970c).

It may be assumed that empirical evidence of the differences between successful and unsuccessful instruction may be accounted for in terms of such basic components of the communication process. The elaboration of these components to the point where they could account for a substantial proportion of the differences between successful and unsuccessful instruction would then be a model or a theory of verbal instruction. This report discusses seven experiments that attempted to examine some aspects of such a theory. The first one examined encoder variability; the next four examined factors involved in the comprehensibility or decodability of the messages, that is, the structure of optional messages; the sixth examined some factors in teachers (encoders) that may be related to their ability to generate effective messages; and the seventh examined the role of feedback in the communication process.

The general methodology most appropriate to this type of analysis involves "communication games," a technique that has been used most extensively in studying the development of children's speech (Flavell, Botkin, Fry, Wright, & Jarvis, 1968; Glucksberg & Krauss, 1967). Usually in communication games two persons, an encoder and a decoder, are separated from each other by a screen; the encoder tries to describe a pattern or an object in such a way that the decoder can select it from a set of alternatives. These tasks can easily be modified and extended for use in studying the teacher-child communication process. The advantages of the technique are twofold: it permits each aspect of the communication process to be studied independently, and more important, it permits much simpler short-term experiments for examining various aspects of the model than are possible in most contemporary teaching research. The assumption made is that these communication games duplicate the essential aspects of information transmission and processing involved in verbal instruction. The relative simplicity of the system may make possible an account of this aspect of teaching that apparently is not forthcoming from direct classroom observation.

EXPERIMENT 1: ENCODER VARIABILITY

In the simplest model of the instruction process (Smith, 1960), the teacher, taking into account the curricular objectives, generates a message that is then transmitted to and decoded by the student. Given a common teaching objective, do teachers differ in the messages they formulate to reach it? If so, in what ways do their messages differ? Do these different messages have a differential effectiveness?

The first two of these questions are specific versions of a more general question, "What do teachers do when they teach?", that has led to the various taxonomies of teacher behavior. The limitations of such a priori taxonomies were pointed out earlier. From our perspective, the only way to define and evaluate the categories generated by the first two questions is to examine them in the light of the third question: that is, the differences between the messages must be examined in the light of the differences in their effectiveness. Hence the first two questions are primarily definitional or conceptual, and the third is empirical (Thomas, personal communication, 1970). What, then, are the consequences of the variability among teachers in the messages they formulate?

Subjects

A heterogeneous group of twelve employees of the Ontario Institute for Studies in Education (OISE), consisting of experimental and clinical psychologists, research staff, students, and secretaries served as encoders in this experiment. Thirty-six first-grade children served as decoders.

Materials

A six-inch square wooden board with four round recessed holes, (d = 2"), each of which could hold a wooden form, was constructed. The wooden forms placed into the form board varied on four binary dimensions: shape (round or triangle); size (large [d = 1-1/4"] or small [d = 3/4"]); color (black or white); design (a star or a cross). Thus optimal messages contained six bits of information, as for example, "Put the small, white triangle with the star on it into the top, left hole of the board."

Procedures

The teachers (encoders), having been shown the forms in position in the form board, were asked to describe the pattern on their board to a hypothetical first-grade child so that he could construct the identical pattern. They were told that their messages would be recorded on tape and played to first-grade children who would be given an identical form board and an appropriate pile of wooden forms. For warm-up purposes, the encoders first described the position of wooden forms in a two-hole form board. Then their messages describing the four-hole form board were tape-recorded, transcribed, and analyzed in terms of length and content.

The two longest and the two shortest messages were subsequently played to groups of first-grade children. Decoder performance was scored in terms of the percentage of correct forms placed in the form board. The effectiveness of the messages was gauged by the decoders' scores.

Results and Discussion

An inspection of the transcripts showed that the messages differed widely in their informational adequacy, prestructuring and rapport building, redundancy, pauses to allow for the child's presumed slower processing, and general "illocutionary force" or persuasiveness. As a summary index, the messages varied greatly in length of recording time, from 17 seconds to 4 minutes and 40 seconds. They also differed widely in their decodability by young children, as measured by the percentage of correct form-board placements made by children listening to the taped messages. One of four messages--the two shortest and the two longest--was played to each of four groups of first-grade children with nine children in each group. Table 1-1 shows that for these extreme cases, decodability was

TABLE 1-1

Message Length and Decodability

Message length[a]	Decodability (% of correct placement by nine children)
17 sec.	36%
1 min. 5 sec.	69
2 min. 25 sec.	78
4 min. 40 sec.	88

[a]For the two longest and the two shortest of the messages generated by twelve encoders.

closely related to message length. It would be naive, however, to conclude that message length alone is the critical factor. There were long messages that were confusing because they failed to make their assumptions explicit, because they omitted an important cue, or the like.

That many factors other than length would be involved in determining the effectiveness of these messages is obvious from even a cursory examination of two transcripts, the first being one of the shortest messages and the second being the longest:

> In the lower right-hand corner put a triangle. Upper right-hand, a circle, large, black. Upper left-hand corner, a small white triangle, and in the bottom corner, bottom left-hand corner, a small white circle.
>
> Now you've got a little square board in front of you . . . and you see some little holes with little blocks in them. . . . All right, now. Now you see some of the blocks are black, some are white and the white ones are smaller than the black ones, and all the little blocks fit into little round places in the board. . . . Now I am going to ask you to take the blocks and put them into the little holes in the board, the little round holes. Don't start yet; I'll tell you about each one, one at a time. Are you ready? . . . Now let's take the round black block. Have you got it? The round black block. All right? Now which is the top of the board? Do you know where the top of the board is? There are two holes at the top. . . . And by now maybe you know which your right hand is. All right? Now, see the two holes at the top there. I'd like to see you take that round black block and put it into the right hole, at the top. Have you got that? Now, again at the top you have an empty hole there on the left side, right beside the round black one you have an empty hole at the top and we are going to put another block in there. . . . Now let's find one of the little white ones. . . . Do you know what a triangle is? You've got two little white ones, haven't you? A round one and one with three corners on it that we call a triangle. Now let's take the little white one, the triangle, and put it in the other top hole on your left side. . . . Have you got it? Now we've still got two holes at the bottom, haven't we? We've got two blocks to put in yet. Let's take the big black triangle--that's the one with three corners on it. Did you find that? The big black one. Now we've got two holes on the bottom. Let's put the black triangle in the right-hand hole at the bottom. Now, if you have done that right, you will have the black triangle under the black circle, just underneath it . . . in the hole below the black circular, the black circular block. Now, that just leaves us with one hole, doesn't it? And you see what piece you have left. It's a little white circle, and that goes in the last hole. That's on the left side at the bottom. Put it in there and we are all finished.

These transcripts reveal both the informational inadequacy of the first message--the child has no way of knowing which triangle to pick up first--and the marked sensitivity of the second message to the child's requirements. They also illustrate the extreme difficulty of identifying directly from the transcripts the qualities that make a message effective. For this reason we have tended to rely on decoder performance as the criterion for judging the effectiveness of encoding.

Some limitations in this type of study could be reduced in further research. First, the use of a hypothetical audience rather than a live audi-

ence appeared to introduce irrelevant variance into the results. Second, it is possible to analyze more fully the messages as well as their decodability by both children and adults. It would be interesting, for example, to see whether adult listeners could decode the short messages as easily as the long. In the present case it was decided to postpone this intensive analysis until more ecologically valid transcriptions could be obtained in subsequent experiments.

This preliminary study does indicate that adults vary substantially in their style and effectiveness in encoding even simple messages for young children. It also indicates the extreme difficulty of specifying what accounts for that effectiveness. The succeeding experiments were designed to investigate a number of the factors that influence the decodability of messages.

EXPERIMENT 2: THE DEVELOPMENT OF OPTIMAL MESSAGES

The preceding study showed that teachers (encoders) vary in the messages they formulate, and that some messages are more effective than others. In this study an attempt was made to isolate factors that may explain why some messages are more effective than others. The factors considered here are those attributable to the verbal messages themselves, that is, factors assumed to be constant for all Ss rather than pertaining only to specific encoders or decoders. While such factors overlap somewhat with those listed by Gagné and Rohwer (1969), mentioned earlier, they are more specifically relevant to the decodability of verbal messages. They fall into four categories.

First, there are factors related to the comprehensibility of an utterance. It is easier to extract information from grammatically simple sentences, for example, than from those involving such transformations as the passive and negative (Gough, 1965; Slobin, 1966). It is easier to extract information from sentences containing the unmarked form of an adjective (big, more) rather than the marked form (small, less) (Clark & Chase, 1970; Donaldson & Wales, 1970); from sentences that are contextually appropriate (Huttenlocher, 1968); from pairs of sentences that share the same underlying grammatical structure (Clark, 1969); and from sentences that have the same grammatical structure as the questions asked about them (Wright, 1969). Further, Morton (1969) has similarly shown that contextual factors, such as preceding sentences, can increase the ease of word recognition.

Second, these are factors related to whether or not the information extracted from a sentence is later recalled. Some evidence indicates, for example, that both the grammatical structure and the underlying logical structure of messages determine what and how much information is recalled. Wanner (1968) found that the probability of recall of a noun subject was determined not simply by the number of times that noun occurred in the surface structure, the actual spoken sentence, but by the number of times it occurred in the underlying or inferred deep structure; that is, the way Ss process sentences determines what they will recall. Frase (1969) has reported on an interesting set of experiments showing that whether the constituent information units in a paragraph will be recalled depends on their relevance or utility in verifying the proposition the reader was considering. Constituents irrelevant to that proposition tended to be ignored or forgotten. Other factors relevant to the recall of verbal information, including frequency, serial position effect, meaningfulness, imagery, and the spacing of practice sessions, are discussed in the verbal-learning literature reviewed by Gagné and Rohwer (1969).

A third set of factors pertains to the type of achievement being assessed. In the simplest case, one form of instruction may be more effective for short-term recall, and another for long-term recall; still another may be more effective when the criterion is transfer rather than

recall. To elaborate a point suggested previously, whereas in recall tasks pictures have consistently been found a superior means of instruction to words (Paivio, 1969; Rohwer, 1970); in transfer tasks, words have been found superior to pictures (Johnson, 1969). Olson (1970c) has suggested that the ambiguity of a picture (the large number of alternative codings) may be an aid to recall, whereas the unambiguity of a word (the small number of alternative codings) may be an aid in transfer tasks.

The fourth type of factor involves the informational requirements of the task. A message appropriate for one problem may not be adequate for another. For simple problems of the type used in this experiment, a direct measurement of the information necessary is possible (Miller, 1956; Shannon & Weaver, 1949); for more complex problems, such a measurement is more difficult, if not impossible. A related problem is that different listeners may be entertaining different sets of alternatives and therefore require different information. We shall return later to the interaction of decoder abilities and expectations with various verbal instructions. The aim of the present study is to demonstrate that some of the factors determining the effectiveness of verbal instruction are empirically accessible and to indicate how they may be incorporated into a theory of instruction. This experiment examined how certain alterations to a basic message--specifically, the use of prestructuring, repetition, and pauses--affected the listeners' ability to recall the message.

Subjects

The subjects for this experiment and Experiment 3 were 60 Toronto first-grade children, randomly assigned to three groups of 20 each. Two of these groups were used in Experiment 2, and the remaining group in Experiment 3.

Messages

The basic message was the simplest message that was judged informationally adequate for its purpose, recorded at conversational tempo: for example, "Put the small white triangle with the star on it into the top, left space" (message length, 10 seconds). The elaborated message was made from the basic message by splicing into the tape recording three types of elaboration: (1) prestructuring, in which Ss were told the essential form of the message but were not given any particular information; (2) repetition of each half of the message; and (3) pauses (one after the first four bits of information, followed by one after every succeeding two bits) to allow for processing by the children. An example of an elaborated message containing all three types of elaboration follows:

> Look at all the pieces of wood on the table. There are black ones and white ones, circles and triangles, small ones and large ones. Here is a circle; there is a triangle. Now look at the board. There are places to put things in it. I'm going to ask you to put the circles and triangles in certain places.

> Put the small white triangle . . . [3 sec.] . . . the small white triangle . . . [3 sec.] . . . in the lower left space . . . [3 sec.] . . . the lower left space . . . [10 sec.] . . . .

There were numerous illustrations of all three forms of elaboration in the original transcripts obtained from adults in Experiment 1.

The proposed strategy was first to compare the basic message to one with all three types of elaboration (fully elaborated message). Then, if clear effects were found, the three types of elaboration would be examined separately. In that event, however, the effects of the combined types of elaboration were such that the experiments on each type individually were not carried out with these messages.

Procedures

A crayon, a cup, a coin, a box, a pencil, and scissors were put out on the table. The subject entered and was seated facing E. Four warm-up problems were given--two by E and two on tape--using these simple materials. Any errors made were pointed out to the S after each problem, and then the problem was repeated. After all the warm-up tests were finished, the objects were cleared from the table.

The form board and wooden pieces were then uncovered. To Ss in the basic message group, E said: "You see all these pieces here? The tape will tell you what to do. You will have to remember because the pieces will be covered up. Remember which piece and where to put it." To Ss in the elaborated message group, E first said: "You see all these things here? The tape will tell you about them. Listen." Then the first part of the tape, the prestructuring or task familiarization, was played. Only after this preparation were the elaborated message Ss told: "The tape will tell you what to do. You have to remember because the pieces will be covered up. Remember which piece and where to put it."

For both groups the pieces and the board were then covered and the first problem played from the tape recorder. As soon as the message was completed, the materials were uncovered and the child chose the form and placed it on one of the four spaces (form and placement were recorded by E). Before each of the remaining three problems, the S was reminded to remember the piece and where it should go. The pieces were covered and uncovered as in the first problem.

Results and Discussion

The children's mean performances on the basic message and the fully elaborated message were almost identical, 70 percent correct as compared to 71 percent correct. Thus the three forms of elaboration--structuring, repetition, and pauses--were apparently unimportant to the decodability of the message. On closer examination, however, it becomes clear that the error patterns for the two groups were considerably different (see Figure

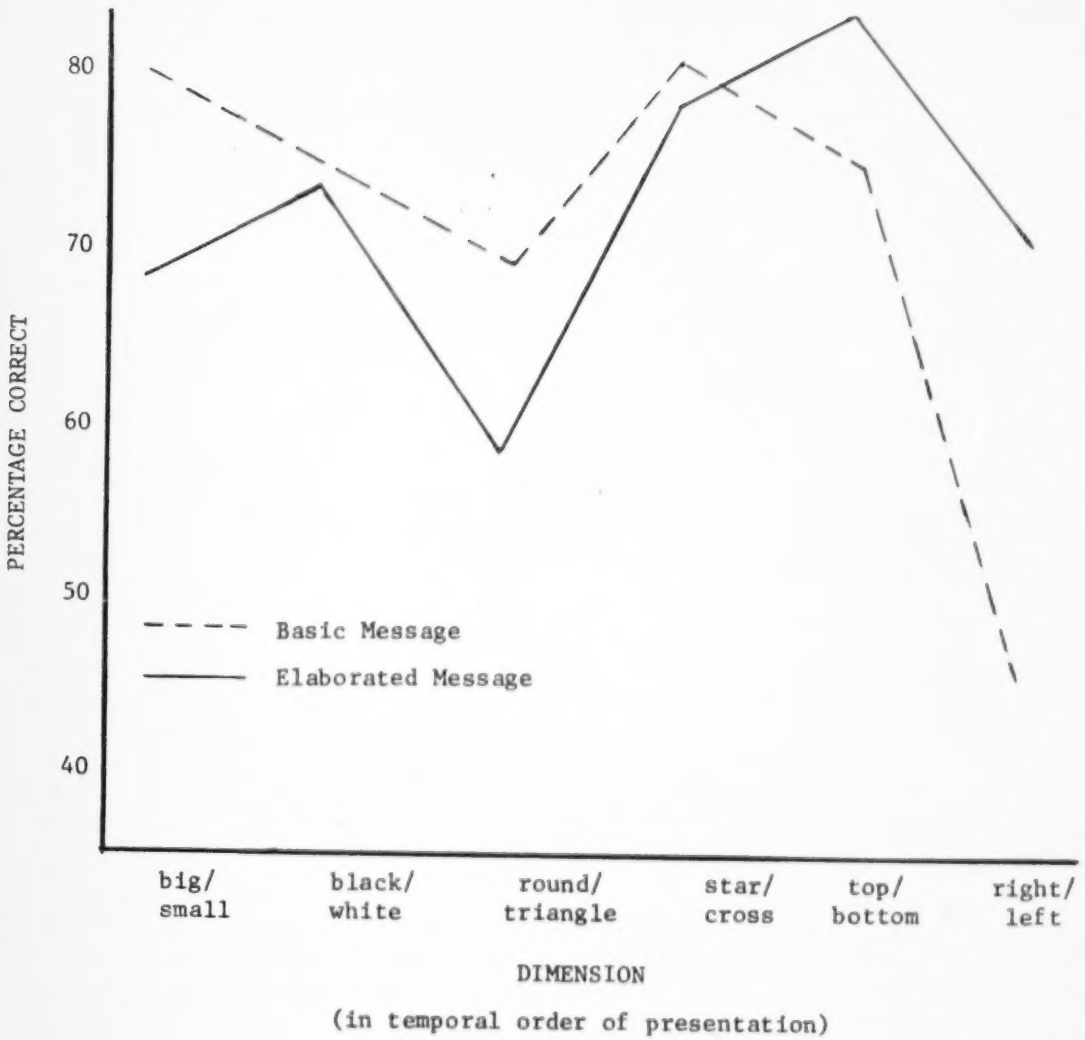

Fig. 2-1. Percentage of binary features correct over four problems for the basic and elaborated messages. N = 20/group.

2-1); that is, though the overall error rates (70 and 71 percent) were almost the same, they were produced by quite different performances by the two groups of children. The factors in the elaborated message influenced which bits of information were decoded, though not the total amount of information decoded. Table 2 presents this analysis of variance. It shows that whereas there is no significant difference between the messages, there is a significant difference between bits of information (dimensions) within the messages and interaction between these message components and the message type. The nature of this interaction is illustrated in Figure 2-1.

TABLE 2-1

Analysis of Variance for the Dimensions Correctly Decoded for the Basic and Elaborated Messages

Source of variation	SS	df	MS	F
Between subjects	55.47	39		
A (message)	.11	1	.11	.08
Ss within groups	55.36	38	1.46	
Within subjects	229.83	200		
B (dimension)	25.93	5	5.19	5.35**
AB	19.21	5	3.84	3.96**
B x Ss within groups	184.69	190	.97	

**$p < .01$

The first three bits of information were decoded better in the basic message; the fourth was decided at about the same rate in both messages; and the last two were decoded better in the fully elaborated message.

There are several possible explanations for the failure of the elaborations to change the overall effectiveness of the message. The simplest is that the effects of the elaborations essentially canceled each other out. This explanation seems unlikely, however, since the elaborations had a differential effect on the various items. A more plausible explanation is that the effects of the elaborations were canceled out by other conditions in the experiment. Specifically, the information the children were required to store and recall was in the form of independent, binary bits that could not be recoded into larger information "chunks" (Miller, 1956). Consequently the amount of information decoded was ultimately a function of short-term memory span. Since young children can store and retrieve relatively little information (4 to 6 digits), a bit of information inserted into a subject's memory span by the elaborations used in the study entered at the expense of one of the bits already inserted by the basic message. Hence the number of bits decoded remained constant, while the elaborations in the message determined which bits were retained.

If in such a case the child's short-term memory limits his decoding of the message, the teacher's ability to transmit the message is obviously restricted. Presumably there is nothing the teacher can do to increase the number of bits processed and recalled by the child. But a good teacher

presents the information in such a way that the child's limited memory span is not exceeded. For example, he may break the message into short components, letting the child process one component before proceeding to the next, or he may adjust his speaking tempo to the processing requirements of the children. This is what, in fact, the majority of the adult encoders in Experiment 1 have done, primarily through their use of pauses. Such techniques were investigated further in Experiment 3.

EXPERIMENT 3: SOME OTHER ASPECTS OF OPTIMAL MESSAGES

As suggested above in the discussion of Experiment 2, the provisions made for the decoders to process and store the transmitted information may be another important factor in the effectiveness of a message. Thus if the children can make use of some of the information either as or before subsequent information comes in, there is less risk of exceeding their short-term memory span limitations. In Experiment 3 we have examined this proposition first by leaving the wooden forms uncovered while the decoder listened to the message, and second by altering the tempo of the message.

Subjects

The remaining group of 20 first-grade Ss who had been randomly assigned to a group at the time of Experiment 2 were used in the part of the experiment where the wooden forms were on display while the taped message was being heard. Some 24 additional first-grade Ss, divided randomly into two groups, served in the part where the tempo of the message was altered.

Materials

The messages for the first part of this study were identical to the elaborated messages used in Experiment 2. For the second part, two new basic messages were constructed. Both contained the same five bits of information, but the former was spoken at a normal conversational tempo and the latter at a slower tempo (called the "instructional" tempo) that E intuitively judged appropriate for teaching young children. The conversational-tempo message was 54 seconds long, and the instructional-tempo message, 66 seconds.

Procedures

The procedure used in the first part of this experiment was identical to that used with the elaborated message groups in Experiment 2, both in the warm-up problems and in the tape-recorded messages. The only modification was that the wooden forms were left uncovered while the child listened to the message telling him where to place them. In the second part of the experiment the procedure was again generally the same, except that one group heard the message in an instructional tempo and others heard it in a conversational tempo.

Results and Discussion

As predicted, permitting subjects to view the wooden forms while they listened to the message increased the decodability of the message. With the forms uncovered, the children's performance improved to 89 percent correct (as opposed to 71 percent in Experiment 2, when the forms were

covered). The summary of the analysis of variance of these results is presented in Table 3-1; it shows that these differences are highly significant.

TABLE 3-1

Analysis of Variance for Decoding by Twenty Children of Messages with Covered and Uncovered Forms

Source of variation	SS	df	MS	F
Between methods	3222.1	1	3222.1	28.26**
Error	4333.7	38	114.0	

**$p < .01$

Evidently the visible array combined with the elaborated message permitted the child to listen to the first component of the message, which was within his memory span, process this information, and visually select the corresponding form. This operation had the effect of emptying the short-term memory storage space so that the child could then listen to the second component of the message and decode it. The net effect was a substantial increase in the children's success in decoding the message.

The effect of different speech tempos on the decodability of the message was also significant. Whereas the performance of the conversational-tempo group was 76 percent correct, that of the instructional-tempo group was 87 percent correct ($t = 1.99$, $df = 22$, $.05 < p < .10$). Hence we may conclude that if information is presented at a slow enough tempo, processing can occur while new information is coming in, even though there are no pauses to allow for that processing.

It also appears that one can intuitively judge how information should be presented to a child. E formulated no particular plan for transmitting the message to the child, but merely bore in mind that the child would be required to comprehend and respond to the message; that thought in itself was sufficient to make E speak more slowly. Of course, when one is attempting to transmit complicated information, one naturally speaks more slowly than usual. The way a speaker automatically adapts his speech to the nature of the message and the requirements of the listener may help account for the fact that the pauses in Experiment 2 had little effect. Perhaps

the basic message was slow enough to permit the kind of processing required, thus making additional pauses unnecessary.

These preliminary experiments on factors that determine the effectiveness of instructional messages, while showing that the problem is amenable to empirical investigation, have not produced any findings that go substantially beyond what one would intuitively expect. One limiting factor in this work has been the task used in the communication experiments. Although the type of information transmitted has the advantage of being binary and hence measurable in terms of information theory, it has the disadvantage of offering little opportunity for an imaginative teacher to experiment with different methods of presentation. In a subsequent experiment we shall consider the effects of some encoding strategies that are appropriate to more complex material (Experiment 6).

Another issue raised by these studies is the relationship between factors affecting learning and factors affecting teaching. For example, if repetition facilitates learning, must the teacher be repetitious? We shall return to this problem as well (see pp. 83-88).

EXPERIMENT 4: THE EFFECTS OF RECODING AND PUPIL ABILITY ON DECODABILITY

Although Experiment 2 showed that various ways of modifying a verbal message--by prestructuring, repetition, and pauses--seemed to affect the information listeners obtained from the message, the exact relationship between the factors studied and the experimental results remained unclear. The difficulty was apparently caused by the presence of an unexpected additional factor, the children's short-term memory span. As we have seen, the effects of this factor were exacerbated by the fact that the information presented in the experimental message could not be recoded by any obvious means; that is, no structure or organization factors could be introduced. Hence in succeeding experiments the messages were constructed from slightly more complex information that could be recoded in various ways. The concept of recoding is borrowed from Miller (1956), who used it to account for the effects of organization on human memory. He pointed out that while human memory is limited to about seven items or words, plus or minus two, people regularly do such things as process and generate twenty-word sentences. To account for this phenomenon, Miller introduced the concept of "chunking," or recoding, and showed that the number of items recalled could be vastly increased if the items were organized into "chunks": a person could recall about seven chunks, each containing up to about seven items. Mandler (1967) has tested this notion in verbal learning. If adults are verbally given a list of words, they can recall about five; if they recode or categorize the words, they can recall about five from each category, up to about five categories. Parallel recoding systems are involved in the recall of visual material. A child who can recode a picture in a large chunk, such as the concept "war," can presumably recall much more than a child who simply codes part of the picture as "gun." The range of possible types of recodings and their effects is an important area of instructional research.

Experiment 4 consisted of two parts. The first involved the teaching of a sixteen-item word list in seven different ways: a basic message and six different recodings using petition, mnemonics, conceptual grouping, visual grouping, and two combinations of factors. The second involved the short-term recall of a set of discrete instructions. This task resembled digit-span tests, but it involved objects and positions rather than numbers. The instructions were given in three different ways: a basic message and two different recodings using grouping and petition. The first question to be investigated, then, was which recoding strategy or form of instruction was the most effective.

The second question investigated in this study was whether some forms of instruction were more appropriate for certain types of learners. That is, all instruction makes some assumptions about the learner; the question is whether different instructions make different assumptions. It was expected that some aptitude differences among the listeners would interact differentially with some of the forms of instruction used in the experiment. This subject has been discussed by Cronbach and Snow (1969) in

terms of aptitude-treatment interaction (ATI). In their report, which is primarily a thorough review of previous research on this topic, their conviction that such interactions exist and are socially important is countered by their finding that few such interactions, if any, occur in the empirical literature. On the one hand they say:

> Adaption to the individual has been an educational slogan of many schools. Our point is that such adaption has never been systematic because no one has known the principles that govern the matching of learner and instructional environment [p. 13].

But on the other hand they say:

> The most basic conclusion from this literature is that simple characterization of aptitudes and treatments in such terms as "spatial" is unlikely to identify combinations of variables worth investigating. There is no instance where an ATI study defined in terms such as these, using familiar "content" constructs from the PMA, DAT, Guilford, or other such aptitude collections, has led to convincing evidence of interaction. The better-controlled studies of this type have led to convincing evidence of absence of interaction [p. 184].

Part of the difficulty, as Cronbach and Snow acknowledge, is that there is no psychological concept of "treatment," and even the concept of aptitude may have to be revised to develop an understanding of ATI. If treatments are thought of as forms of instruction, it is obvious that there is no theory to account for their range and effects. Of course, describing a treatment is not the same as offering a theoretical account of it. The first may be done in terms of the requirements of ordinary conversation, as, for example, "I told him how to get there," or "I drew him a map," whereas the latter would require an analysis of the information conveyed by the different forms and the processes involved in extracting the information. No such theoretical account exists in the psychological literature. As pointed out elsewhere, "The language of psychology may have an adequate vocabulary for discussing learning, but it has an obvious lacuna of the vocabulary for even discussing instruction. The solution of simply calling each step of the training 'a stimulus' is to beg the question [Olson, 1970a, p. 143]." Other solutions, such as using quasi-theoretical distinctions like "spatial/verbal" and "expository/discovery" without specifying the kinds and amounts of information involved and the kinds of processing required, are also likely to have limited value in explaining instructional effects. Maccoby (1968) has expressed similar reservations about the current training experiments in child development literature.

An important theoretical approach to the aptitude-treatment interaction has been made by Hunt (1966, 1970, 1971) and by Hunt and Joyce (1967) in the concept of "matching models." For Hunt and Joyce an educational approach or form of instruction must be assessed in terms of its appropriateness to the particular learner. A match occurs when the form of instruction

combined with the learner's characteristics leads to the instructional goal. Two recent studies from this point of view (McLachlan, 1969; Tomlinson, 1969) have produced some evidence that the "conceptual level" (CL) of the learner interacts with the form of instruction, the highly structured, explicit lecture approach being more appropriate to the low CL learners and the less structured discovery lesson more appropriate to the high CL learners.

In the present study, Experiment 4, we have examined the possibility of such an interaction between the form of instruction and one aspect of the learner's ability. Specifically, we have examined how aptitude, as measured by a digit-span test, related to children's performance under various treatment (instructional conditions) to determine whether some treatments relied more on this aptitude than others. Digit span was chosen as an aptitude variable because the instructional task with which it was expected to interact was a relatively simple recall task. It was hypothesized that forms of instruction having little structure would be more dependent on memory ability than forms having a high degree of structure.

Subjects

The subjects were 140 fourth- and fifth-grade students from two suburban Toronto schools--the Humewood Public School and the H. J. Alexander Jr. Public School. Their ages ranged from nine to eleven, and there was approximately the same number of boys as girls. The Ss were randomly assigned to groups in the following way. Teachers prepared class lists broken into groups of four so that all the groups were roughly comparable in average intellegence. In the Alexander Public School one of these groups of four was randomly selected from each of two classrooms, thus forming groups of eight. In the Humewood Public School, one of these groups of four was randomly selected from each of three classrooms, thus forming groups of twelve. By randomly pairing these larger groups from the two schools, seven primary experimental groups, each with twenty subjects, produced. For the second set of instructional tasks, these two sections of the seven primary groups were cross-classified to yield three new groups of 43, 49, and 48 respectively. All Ss took the digit-span test in the third phase of the experiment.

Materials

The materials used in this experiment consisted of seven instructional forms for the word list, three instructional forms for the memory-for-discrete-instruction's test, and the digit-span test. These materials were presented by means of tape-recorded instructions and a student answer booklet.

Word-List Lessons. The seven different encoding conditions for presenting the sixteen-item word list were produced by means of five tapes and two printed word lists that were used simultaneously with two of the

tapes. Each tape began with the following instructions:

> For the first part of this test, you are to learn a list of words. Now this isn't very hard. You just have to listen very carefully to the list of words and try to remember them. Later, you will be asked to write the words down, so do listen closely. Okay? Are you ready? Here are the words.

Each tape ended with: "Now try to remember those words."

For the baseline condition, the word list was read through three times in the following random order (time = 90 seconds):

> bank, hotel, hill, car, cabin, beach, pond, horse, bread, ham, cliff, potato, coffee, bus, plane, palace.

The repetition condition was identical to the baseline condition except that the word list was read through six times, and therefore took twice as long (time = 180 seconds).

In the mnemonics condition the words were divided into four groups of four words all beginning with the same letter, as follows:

> 'B': bus, bread, beach, bank
>
> 'C': car, coffee, cliff, cabin
>
> 'H': horse, ham, hill, hotel
>
> 'P': plane, potato, pond, palace.

For each group the common first letter was given ("The first group of words starts with the letter 'B.'"); then the group was repeated three times (time = 90 seconds).

For the categorization condition the words were again divided into four groups of four words, but this time on the basis of conceptual categories (transportation, food, geographical terms, and buildings), as follows:

> bus, car, horse, plane; bread, coffee, ham, potato; beach, cliff, hill, pond; bank, cabin, hotel, palace.

For each group the category label was given ("The second group is names of things to eat."); then the group of words was repeated three times before going on to the next group (time = 90 seconds).

For the categorization tape and visual list condition the categorization tape was played and Ss were simultaneously able to look at the words from a sheet of paper in front of them. The words were arranged to correspond to the categorized order.

The categorization tape and visual blocks condition was identical to the preceding one, except that the visual word list was arranged into four spatially separated blocks, one in each of the four corners of the sheet, one block for each of the categories, as in the categorization condition above.

In the narrative condition the words were put together in a story. The story was read twice. The first time, the words to be remembered were repeated at the end of each sentence, and the second time, the words to be remembered were said more loudly than the other words (time = 195 seconds).

The story follows:

> Once upon a time there was a rich man who lived in a palace on top of a hill. He was so wealthy because he owned a beautiful hotel with a lovely beach and pond. People came by car, bus, or plane from all over the world to stay at his resort. One day he was out riding on his horse and fell over a cliff. Fortunately, he was all right except for a sprained ankle, and he managed to get to a small cabin before it was dark. He asked the farmer for a simple dinner of ham, potato, bread, and coffee, which the farmer's wife gave him. He was so grateful to these kind people that as soon as he was better, he want to his bank and gave them a reward of $100.

The students' answers were scored in the same way for all seven encoding conditions: one point was given for each word correctly recalled, yielding a maximum score of 16.

Memory-for-Discrete-Instructions Test (MDI). This test, developed by Ross Traub, consisted of taped instructions which the listeners were to follow in marking various objects and positions in a provided answer booklet. There were three different taped encoding conditions in the test. The general procedure was as follows: (a) S heard the instructions relevant to a page of the answer booklet, and then was told to begin; (b) sufficient time elapsed on the tape so that S could carry out the instructions (from 15 seconds for four items to 30 seconds for seven items); and (c) S was instructed to turn to the next page. The tapes in all three conditions began with the following instructions: "You will now be told to draw different things on the pages in the booklet in front of you. Listen carefully and do as you are told after you hear the word 'begin.'" They ended with: "This is the end of this part of the test. Please put your pencils down."

The instructions for the last test item in each condition are given below to illustrate how the conditions varied.

Baseline: "Put a cross through the plane, a line under the house, a circle around the car, a cross through the truck, a line under the barn, a circle around the boy, and a cross through the ship. Begin." (Total time for tape = 7 minutes, 15 seconds.)

Grouping: "Put a cross through the plane, the truck, and the ship; a line under the house and the barn; and a circle around the car and the boy. Begin." (Total time for tape = 7 minutes, 15 seconds.)

Repetition: "Put a cross through the plane, a line under the house, a circle around the car, a cross through the truck, a line under the barn, a circle around the boy, and a cross through the ship; that is, cross, line, circle, cross, line, circle, cross. Begin." (Total time for tape = 9 minutes.)

Scoring. To calculate the MDI score for all three encoding conditions, one point was given for each item correctly marked yielding a maximum score of 45.

Digit-Span Test (DS). This test consisted of a series of tape-recorded digits that Ss were to listen to and then write in the provided answer booklet. The same form of the test was given to all Ss. The tape started with the following instructions: "For this part of the test, you will hear a series of numbers. When you hear the word 'begin,' you are to try to write down the numbers just as you heard them called out. Only put one number in a box. Turn to the first page with a row of little boxes and we will start."

The first two items, each four digits long, were practice items; the test items increased in length from five up to seven digits. The digits were read at a measured pace, and sufficient time was left after each series to allow the Ss to write them down--from 15 seconds for four digits to 30 seconds for seven digits. (Total time for tape = 8 minutes.) In scoring the test, one point was given for each number that was both correct and at the correct place in the series. The maximum score possible was 50 points.

The Answer Booklet. Ss were not allowed to use pencil and paper while listening to the word-list lessons. For the memory-for-discrete-instructions and digit-span tests, however, each S was given a 22-page answer booklet. The first page had places for such information as the name, age, grade, school, and sex of S. The next ten pages had pictures of objects such as a house, plane, boat, girl, and so on, which S had to underline, circle, or put a cross through, according to the MDI test. The first two pages had four pictures each, and were used as practice items to make sure that the Ss knew what they were to do. The remaining eight pages had an increasing number of pictures per page, from four to seven, and were the actual test items.

The next ten pages were used in the digit-span test. Each page had a row of seven little boxes on it, a place for up to seven digits. Again, the first two pages were used for practice items, and the remaining eight were for the test items.

The last page in the booklet was used in the recall of the word list and contained a place for each of the 16 words the Ss heard at the beginning of the test procedure.

Procedures

Students were tested in groups of eight or twelve at a time. E first told each group that they were going to be given three short tests to determine their ability to remember. Ss were assured that none of the tests was very difficult and that the results would not affect their grade in school (a source of concern for quite a few).

E then explained that for the first test, the Ss just had to listen quietly to a taped lesson teaching them a list of words. The appropriate tape was then played, depending on the encoding condition of that group. After playing the first lesson E gave each S the answer booklet for the two remaining tests.

E then briefly explained that for the second test, S was to listen carefully to the taped instructions, and then carry out the instructions as well as he could remember. The tape appropriate to the particular MDI condition was then played.

After the second test, E told the Ss that they would be hearing groups of numbers for the last test and would be asked to write them down in the correct order after hearing them. The digit-span test tape was then played.

Following the digit-span test, Ss were told:

> This is the end of this part of the test. Now please turn to the last page in the booklet. There are places on this page for the words you learned at the very beginning. Please write down as many of the words as you can remember. Do not worry about the order or spelling of the words; just write clearly. Okay? Now begin.

They were allowed five minutes to recall and write down the word list.

Results

Word-list recall. Figure 4-1 shows that the different ways in which the word lists were encoded did produce different results: the mean number of words recalled ranged from 3.2 words in the mnemonic condition to 9.05 in the categorization-plus-visual-list condition. A one-way analysis of variance performed on the number of words recalled showed that these differences are highly significant (see Table 4-1). When possible differences among the groups in ability as measured by the digit-span test were statistically controlled by means of the analysis of covariance, the dif-

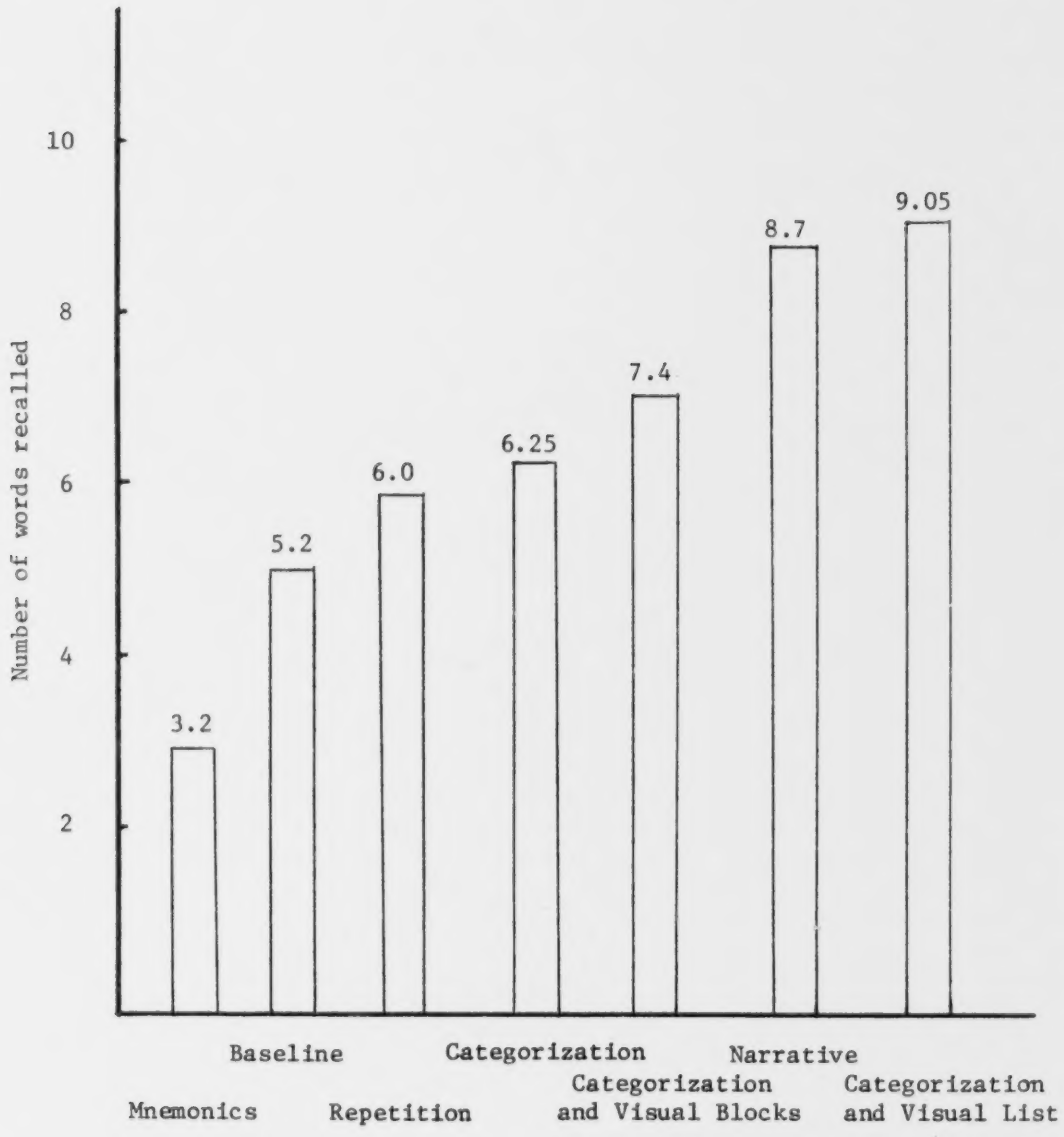

Fig. 4-1. Mean number of words recalled as a function of encoding condition. N = 20.

TABLE 4-1

One-Way Analysis of Variance of the Number of Words Recalled for the Seven Treatment Groups

Source of variation	SS	df	MS	F
Between groups	500.643	6	83.440	14.52**
Within groups	764.100	133	5.745	
Total	1264.743	139		

**p < .01

ferences among the groups in the number of words recalled remained highly significant ($F = 14.42$, $df = 6/132$, $p < .01$). Scheffé comparisons were carried out to determine which of these treatment pairs differed significantly. The results of this analysis are presented in Table 4-2. Any differences of 2.40 or more between groups were significant with $p < .01$.

TABLE 4-2

Comparison of Word-List Recall after Seven Forms of Instruction (Scheffé Method)

Form of instruction	Mean	F 8.7	E 7.4	D 6.25	C 6.0	B 5.2	A 3.2	r with DS
G Categorization and visual list	9.05	.35	1.65	2.80**	3.05**	3.85**	5.85**	.33
F Narrative	8.7		1.3	2.45**	2.7**	3.5**	5.5**	-.13
E Categorization and visual blocks	7.4			1.15	1.4	2.2*	4.2**	.27
D Categorization	6.25				.25	1.05	3.05**	.01
C Repetition	6.0					.8	2.8**	.24
B Baseline	5.2						2.0	-.11
A Mnemonics	3.2							

*p < .05
**p < .01

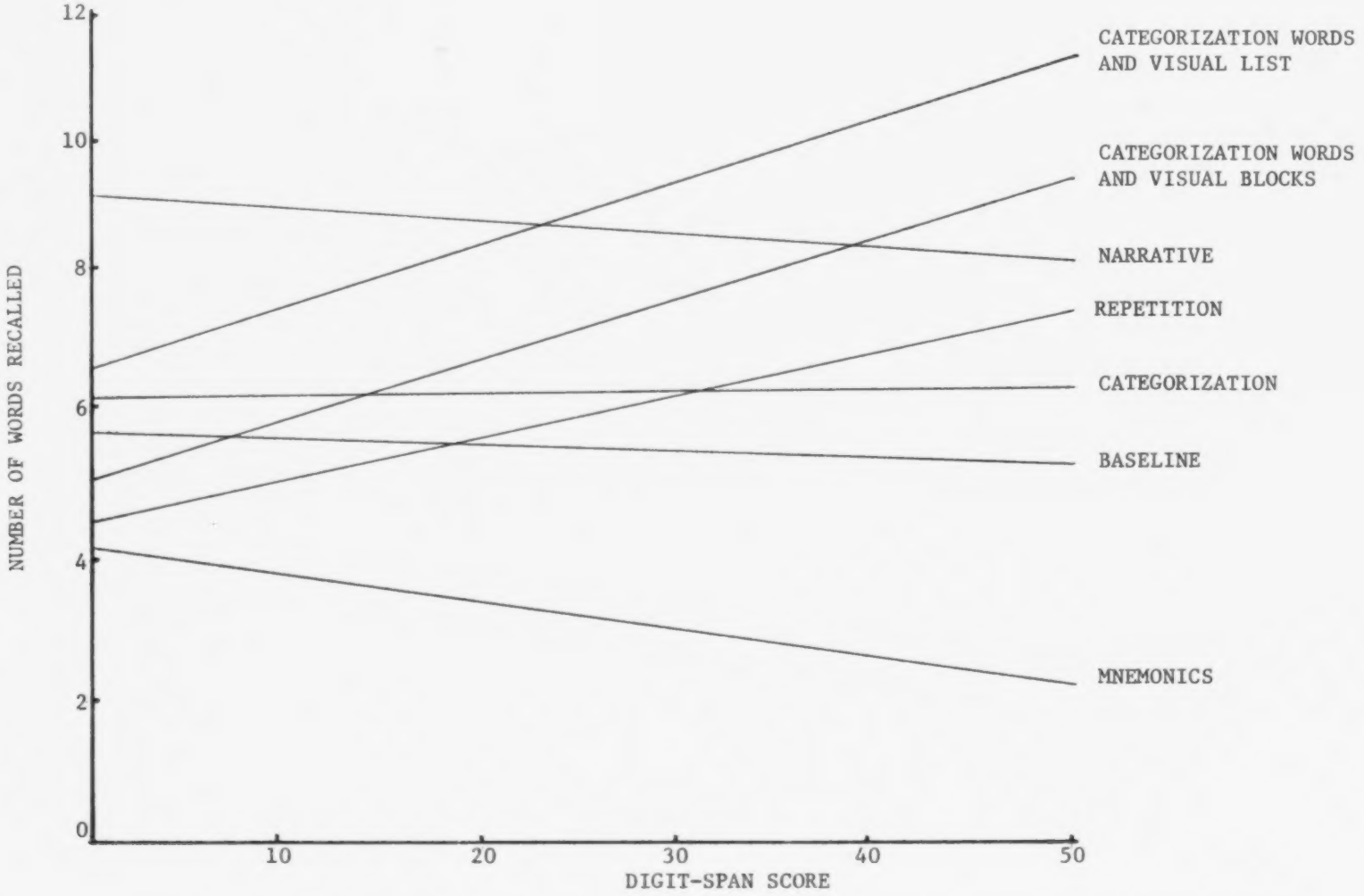

Fig. 4-2. Number of words recalled as a function of digit-span score: Regression analyses.

The table shows that both the narrative condition and the categorization-plus-visual-list condition were significantly more effective than the mnemonics, baseline, repetition, or categorization condition. Furthermore, repetition, categorization, and categorization plus visual blocks were all significantly better than the mnemonics condition.

Aptitude-treatment interaction. Since there were marked differences in effectiveness among the alternative forms of instruction, it seemed possible that some of these forms depended more on the learners' underlying abilities than other forms. Stated in another way, the different forms of instruction might have made differnet assumptions about the learners. To examine this possibility, a regression analysis of recall score and digit-span score (DS) was carried out for each of the seven treatments. These correlations are also presented in Table 4-2. A test showed that none of these values was significantly different from zero. The regression lines for recall score as a function of digit-span score for each of the seven forms of instruction are presented in Figure 4-2. An analysis of the homogeneity of within-class regression showed that the regression slopes were not significantly different ($F = 1.03$, $df = 6/126$, n.s.). That is, there is no evidence of ATI in these data. This finding should be considered tentative, however, in view of the fact that there were only 20 Ss in each of these instructional conditions.

Although the aptitude-treatment interaction was not significant in this study, it is interesting to note in Table 4-2 the apparent tendency for the less successful treatments to correlate negatively with digit-span score and for the more successful ones to correlate positively. This relationship, as assessed by a rank correlation, yields a Spearman = .61, which approaches significance (required value for $N = 7$ is .71). Recall that the more successful treatments were those that had a higher degree of conceptual structure. Thus the relationship between digit-span score and the degree of structure used in presenting the lists is just the opposite of that predicted. While further evidence is needed to establish this point, it appears that in this case good instruction heightened the effects of individual differences rather than diminishing them.

Memory for discrete instructions. Figure 4-3 shows that the different ways of encoding the instructions produced different levels of performance by the children. A one-way analysis of variance of the MDI scores indicated a significant difference among the groups as shown in Table 4-3. When possible differences among the groups in initial ability as measured by the digit-span test were statistically controlled by means

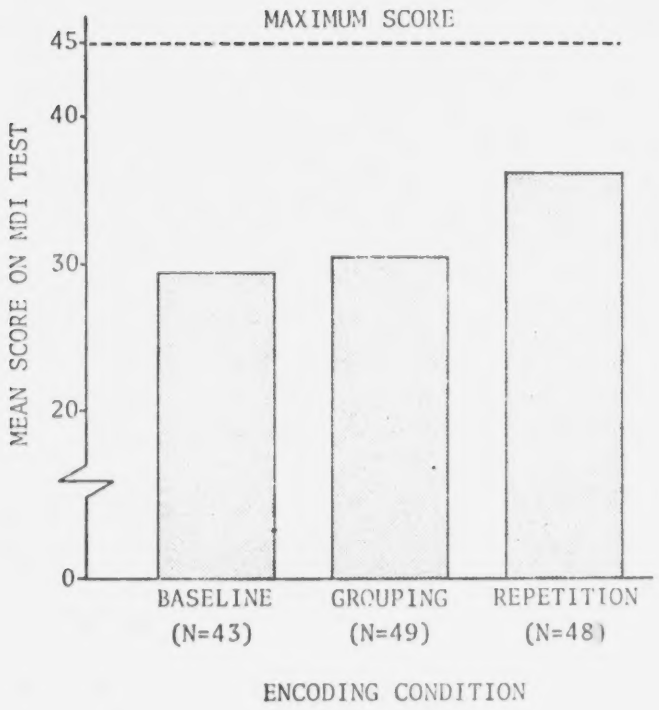

Fig. 4-3. Mean score on the memory-for-discrete-instructions test as a function of encoding conditions.

TABLE 4-3

One-Way Analysis of Variance of Memory-for-Discrete-Instructions Scores under Three Instructional Conditions

Source of variation	SS	df	MS	F
Between groups	1264.44	2	632.22	18.97**
Within groups	4566.79	137	33.33	
Total	5831.22	139		

**p < .01

of the analysis of covariance, the differences among the groups in their performance on this task remained highly significant ($F = 21.44$, $df = 2/136$, $p < .01$). Scheffé comparisons of the differences between the three treatment groups are presented in Table 4-4. Any differences greater than 6.63 were significant with $p < .01$. This table shows that the repetition condition was superior to the baseline condition.

TABLE 4-4

Comparison of Recall after Three Forms of Instruction on Memory for Discrete Instructions (Scheffé Method)

Form of instruction	Mean	B 30.65	C 29.65	r with DS
A Repetition	36.46	5.81	6.81**	.22
B Grouping	30.65		1.00	.34*
C Baseline	29.65			.50**

*$p < .05$
**$p < .01$

As with the word-list instructions, an attempt was made to determine whether forms of instruction for MDI interacted differentially with ability as measured by DS scores. A regression analysis of MDI and DS scores for each of the three treatments yielded the correlations presented in Table 4-4 along with their levels of significance. The results show that DS scores were related to MDI scores under the baseline and grouping conditions but not under the repetition condition. The regression lines for MDI scores as a function of DS scores for each of the three conditions are presented in Figure 4-4. To determine whether there was an aptitude-treatment interaction, that is, whether these slopes are significantly different, a test for the homogeneity of within-class regression was carried out. This analysis showed that the slopes are not significantly different for the three forms of instruction for MDI ($F = 1.01$, $df = 2/136$, n.s.).

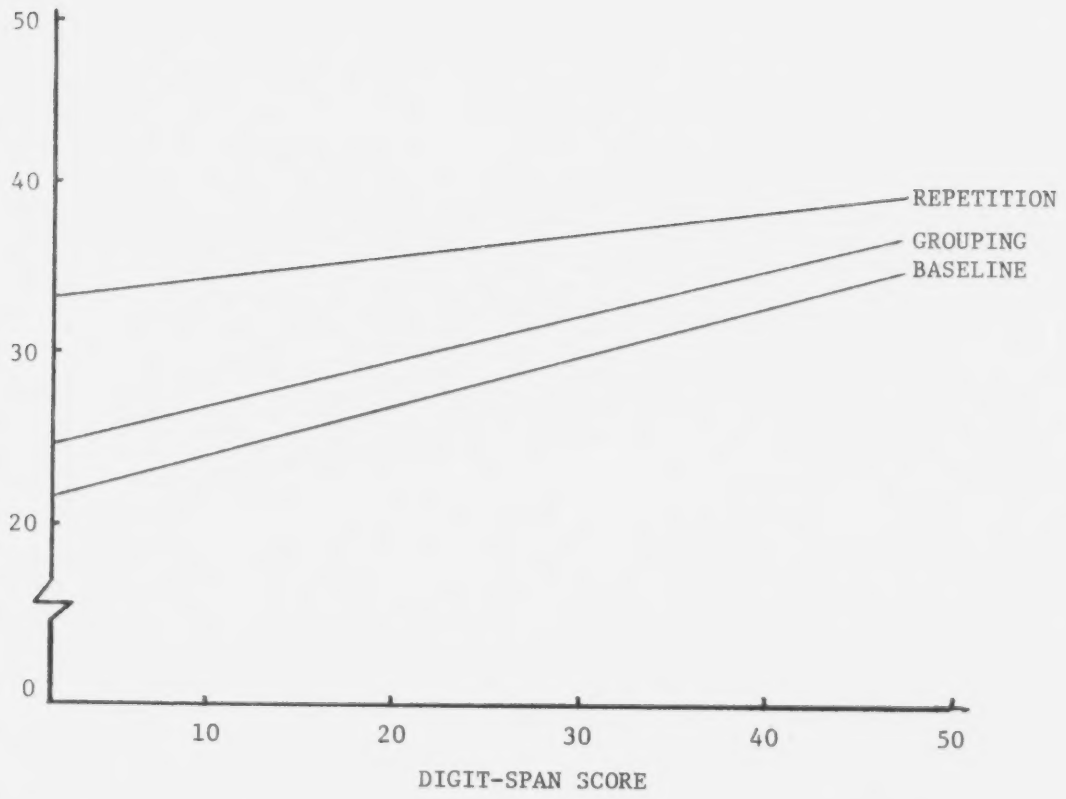

Fig. 4-4. Regression lines for memory-for-discrete-instructions score as a function of digit-span score.

Although a hypothetically relevant aptitude variable, DS failed to interact differentially with the three forms of instruction presented in this experiment, DS ability might be found more relevant under poorer forms of instruction. This pattern, if substantiated, would be in line with the prediction made at the outset.

Discussion

Factors affecting the recall of taught materials. Several inferences may be drawn about the factors that made some messages easier for the children to recall than others. It should be noted, first, that in all treatments Ss were exposed to the information items three times, except in the repetition condition, where they were exposed six times. The narrative condition presumably operated by organizing the list, so that any one component of the story recalled would serve as a cue for the remainder of the story. Different narratives could thus be expected to be differentially effective, depending on their degree of integration or structure.

The same type of organizational factor could be expected to be at work in the categorization condition; the effect of this form of instruction was not statistically significant, however. Perhaps formal semantic categories are not as effective as a simple narrative for organizing information. Perhaps the difference lies in the fact that the narrative forms one structure, whereas the category systems in this experiment formed four independent structures with no link between them. Or perhaps it lies in the fact that while the number of exposures to the information items was the same under both conditions, the narrative message was almost twice as long as the category list, thus giving Ss more time to rehearse and reflect on the material they were being taught.

The visual supplement to the verbal list had an important effect on recall. Although this finding bears out the teacher's intuition that two channels for presenting information are better than one (hence, for example, that important points should be written on the blackboard as well as taught verbally), it is not clear why the visual list was effective. Three possible explanations could be investigated in further research. First, a visual stimulus, being longer in duration than an auditory one, may thereby focus greater attention or concentration on the item. Second, the presence of the visual list may permit a child to rehearse the words more frequently by looking back up the list as he hears each new item. Third, and perhaps most important, the presence of the visual list may enable a child to scan the list for possible groupings within it. If this list explanation is true, there should be greater incidence of clustering in recall for Ss in these experimental groups.

In preparing the visual arrays the experimenters had considered the possibility that putting the word groups into spatially separated visual blocks would add a spatial cue that might aid in the recall of the items. But the visual blocks were less effective, if anything, than the simple visual list. Further work should be done on the effect of such spatial cues on recall. Murdock (1967) has reported that spatial cues facilitate the recall of an auditory message. There is some evidence, too, that readers can estimate with great accuracy the position on the page where various information appeared (Rothkopf, in press). It is not known whether this ability facilitates recall, though the present evidence indicates that it does not.

The poor performance of the Ss in the mnemonics condition suggests that mnemonics is not an effective means of organizing information. This finding may result from the fact that children are not currently encouraged to use such memory devices as much as they were a few decades ago. Possibly a more elaborate mnemonic would have had a better effect; if the four key letters had themselves formed a word, they might have been more useful. Possibly this type of memory device is incompatible with other methods, such as conceptual grouping and petitions, which children may use more spontaneously.

A much clearer picture of what the Ss were doing and why some conditions were better than others could be obtained from an analysis of clus-

tering in recall. If mnemonics were being used, for example, there should have been some tendency for the Ss to recall either all or none of the words beginning with the same letter. If categorization were being used, the Ss could have been expected to recall some categories and not others. This sort of analysis has been successfully used in research on memory (Bousfield, 1953). It was not performed in the present experiment, however.

Neither adding one level of repetition nor categorizing the word list had a significant effect over the baseline condition, though there were marked differences in the expected direction. Although both of these factors are undoubtedly important, they may have failed to yield significant differences because of the activities of the children themselves: that is, the children may have compensated in subtle ways for the inadequacy of the lesson. In another experiment (Exp. 6) it was observed that, given a very minimally presented lesson, the children rehearsed themselves, until they thought they could remember the content. The children in this experiment, knowing what the task called for and finding the message somewhat impoverished, may have rehearsed the words themselves or sought some way of organizing the words in order to remember them. Flavell, Beach, and Chinsky (1966) have found that beginning at about school age children do in fact tend to rehearse word lists to themselves, even if they are not specifically instructed to do so. Moreover, several researchers (Bousfield, 1953; Mandler, 1966; Tulving, 1968) have found that adults tend to organize items into groups and then recall the groups. If the Ss in our experiment rehearsed or organized the words independently of the lesson, their activities interacted with or perhaps compensated for those of the teacher. There is, therefore, a subtle interplay between what the teacher does and what the learner does--not just a cause-effect relationship, as we had assumed at the outset (Smith, 1960), but rather a compensation or cybernetic relationship. That is, the less the teacher does, the more the children do. This proposition holds true only when the children know the goal and actually pursue it.

This interaction makes the analysis of teaching <u>per se</u> extremely difficult. Studies that examine the mean gain in achievement and try to relate it to the performance of a teacher--that is, studies that treat teaching-learning as a cause-effect relationship--are almost certain to fail. It is not what the teacher does but what the children do that is critical. If the children share the teaching objective, they may resort to their own devices to attain it, regardless of what the teacher does. In the context of the present experiment, it becomes apparent that even though, say, repetition may be an important factor in children's learning, the lesson need not be repetitious as long as the children's learning contains an element of repetition. Thus under an instructional condition with no more repetition than the baseline condition but in which the children were told the nature of the task and asked to repeat each word to themselves five times after hearing it, one could reasonably expect better achievement than under the baseline condition and perhaps even the repetition condition.

In regard to the MDI materials, the repetition of the message was an important factor in the decoder's performance. Possibly this condition might be better considered a simplification of the message rather than simple repetition. That is, since the repeated version of the message was much shorter than the original version, it may have achieved its objective chiefly by leaving more room in the children's short-term memory span for the marking instructions. This possibility was not noticed when the task was being designed. As for the grouping condition, its failure to have an effect is difficult to explain. Since the grouping message was the same length as the baseline message, there was no additional time factor at work. Perhaps whatever advantage the grouping method might have gained was simply nullified by the loss of linear order in the items described. In any case, the degree of recall was again a function of the way the messages were taught. But even these effects were undoubtedly closely tied to the activities of the decoders.

Many questions remain about the effects of various recoding strategies. For example, which would be more effective--two different recodings, each presented once, or the same recoding presented twice? Another area for investigation is the appropriateness of various forms of recoding to different objectives. The recodings studied here aimed only at increasing recall; different results might have been obtained if the objective had been a transfer task. In an analysis of the effectiveness of pictures and words as visual means of transmitting information, Olson (1970c) concluded that words were more suited to transfer and concept formation, whereas pictures were more suited to recall and paired-associate learning.

Aptitude-treatment interactions. In this study the significant treatment effects did not interact differentially with the ability of the learners. That is, no particular treatment was found particularly appropriate to one kind of learner as determined by digit-span scores. It is possible that a wider range of aptitude measures might have detected such an interaction.

As pointed out above, however, the interaction between what the teacher does and what the child does is also important in determining the outcome of an instructional form. What the child does in this regard is not well represented by conventional aptitude tests. Rather it is a function of what the child considers the goal to be, his opinion of its value, and his ability to provide an appropriate form of activity to complement that of the teacher in reaching that goal. Tests may indicate whether or not the child had the skills necessary to supplement the teacher's lessons (e.g., conceptually grouping a word list, looking something up in a reference book, reading a diagram), but these skills are of a different order than the abilities measured by conventional aptitude tests. Moreover, tests of such skills would have to be supplemented by specific situational indices of how seriously the child took the task and how well he knew the teaching objective. If a child does not know the objective, he obviously cannot make an independent contribution to reaching it. Snow (1970) has discussed instruction in a related fashion by considering it an "externalized aptitude, or prosthetic device, which compensates for the weakness in a particular kind of learner [p. 77]."

We have argued above that different treatments rely on different skills or aptitudes on the part of the learner. Some children, especially those at certain stages of development, may lack some of these skills, thereby making certain treatments inappropriate for them. To return to the earlier example, a low degree of repetition in the teacher's message may require a high degree of repetition or rehearsal on the part of the child in order to reach the teaching objectives. Young children may not be capable of supplying that degree of repetition, as Flavell, Beach, and Chinsky (1966) would lead us to expect, and hence a message incorporating little repetition would be inappropriate for them. The same point holds for skills of organization or conceptual grouping in tasks involving memory.

At the outset of these experiments we had considered examining the relationship between children's aptitude, in terms of developmental level, and the effectiveness of various forms of instruction. Thus, in addition to examining the effectiveness of the various forms of instruction considered in this experiment at the fourth- and fifth-grade levels, one could examine the effectiveness of these same lessons when presented to younger children, who in Piaget's terms would be preoperational. It is reasonable to expect, for example, that older children would do particularly well at learning word lists by means of conceptual groupings, since they have generally mastered the technique of class inclusion, the logical form for organizing such lists. Younger children, by contrast, not yet skilled in class inclusion, might benefit more from repetition or from a narrative presentation. Although these possibilities have not been examined directly, some related evidence is considered in the next experiment.

EXPERIMENT 5: STAGE OF DEVELOPMENT AS AN APTITUDE VARIABLE IN DETERMINING THE EFFECTIVENESS OF A VERBAL INSTRUCTION

A wide range of evidence indicates that children in our culture develop the ability to perform certain logical operations, among them class inclusion, when they are about six or seven years old. Before this age, Piaget has shown that children's thinking is limited by the specific forms of actual experience. Hence the child may "run off" reality sequences in his head, just as he might have done in his overt action, but he is unable to reverse, reorder, schematize, or refashion events within an integrated cognitive system (Flavell, 1963, p. 158). The development of concrete operational thought is marked by the appearance of several modes of organization, such as the multiplication of relations, seriation, class inclusion, superordinate classification, and cause-effect relations: "The mastery of the inclusion relation, with all that its mastery implies, is the 'sine qua non' of the Concrete Operational stage . . . of logical classification [Flavell, 1963, p. 306]." At this stage of development a child can combine elements to form a superordinate class, retaining the concepts of the class, its elements, and the relationship between them. This ability in turn permits the child to consider the relationship between a class and a subclass.

This point is illustrated by one experiment (Piaget, 1960) in which children were asked to make (or imagine making) necklaces out of twenty wooden beads, seventeen of which were brown and three of which were white. The children were then asked, "Which necklace will be the longer, the wooden or the brown?" The younger, preoperational children replied that the brown would be longer because there were only three white. They did not appear to be able to hold "wooden" and compare a subunit "brown" with the total unit "brown and white."

In another experiment on class inclusion and hierarchical classification (Inhelder & Piaget, 1964), only 46 percent of the five-year-olds could answer correctly that there were more flowers than primulas, whereas 82 percent of the eight-year-olds gave the correct answer. Again the results demonstrated that young children have difficulty understanding the relationship of a conceptual whole to the parts of which it is composed. The evidence is consistent in showing that class inclusion and superordinate relations are developmental and that they tend to appear in children at about seven years of age in our culture.

Now consider how this development may effect verbal memory. One of the favorite strategies of adults for remembering a group of verbal items is to put them into a class-inclusion relationship, with a superordinate term that includes the set of specific items (as in the preceeding instructional experiment). Miller (1956) has proposed that though the immediate memory span may be limited to about seven items, we normally exceed this limit by "chunking"-- organizing several discrete items into a more general item that can then be treated as a single item. Mandler (1967) and Tulving (1968) have both provided considerable evidence that items are in fact

categorized by Ss and that this categorization determines the number of items recalled. But it is the ability to perform precisely this operation that we have found to be developmental. It follows that the ability to benefit from the intrinsic categorical structure of a word list is developmental; that is, it can be expected to help older children remember the list, but is essentially irrelevant to younger children. In regard to instruction, a lesson using superordinate organization would presumably be beneficial to older children but confusing to younger children.

Since the complete details of the method and results of this experiment are reported elsewhere (Olson, Hildyard, & Crosbie, in preparation) from the point of view of the development of children's memory, only a summary statement is offered here.

Subjects

The subjects were 128 children selected from two middle-class schools in Toronto. There were 16 Ss, eight boys and eight girls, from the kindergarten level (mean age = 5-6 years) for each of four experimental conditions, and a corresponding set from the second-grade level (mean age = 7-8 years).

Experimental Lists

The same sixteen-item word list was presented to all Ss by means of pencil drawings of objects corresponding to these words. Each item was on a separate five-by-eight-inch card in a ring binder. The pictures were presented to the Ss in one of two orders:

1. Ordered lists (O). The items in the list were grouped so that they could be categorized under a general heading in a class-inclusion relationship. The four superordinate categories were food, clothes, toys, and animals.
2. Nonordered lists (N-O). These lists were presented in a scrambled order, so that no two items from the same category were in sequence.

Within the limitations of these two ordering principles, the order of the words was randomized in each of the four presentations.

The second factor in the presentation of the word lists was the provision of the superordinate terms for each of the groups of words, as follows:

1. Superordinate given (S). Before going through the list of words each time, the child was told that he was going to see some pictures of food, clothes, furniture, and animals.
2. No Superordinate given (No-S). The categories were not explicitly named.

The ordering and superordinate conditions, along with grade level and sex of the Ss, constituted the "between Ss" factors; the four recall trials

given each child constituted the "within Ss" factor. These five factors, plus a Ss factor that was later collapsed to yield the error term, were the basis of the subsequent analysis (Clifford, 1968).

Procedures

Each S was given four successive free-recall trials on the entire list. Each card was exposed for four seconds. At each exposure the child was to name the picture; if he was unable to, E provided the correct word. At the end of each trial, S was requested to recall as many of the words as possible. After the fourth recall trial, E asked S to try to recall or to generate the superordinate category names.

Results

The mean recall scores were analyzed by means of a six-way analysis of variance, yielding the following significant F scores:

1. The grade of the S. As expected, the second-grade children recalled more words than the kindergarten children.
2. The sex of the S. Among both kindergartners and second graders, the girls were able to recall more words than the boys; this effect was more pronounced at the kindergarten level.
3. The ordering of the words in the list. Ss given the list arranged in homogeneous groups (e.g., the "food" group--hamburger, apple, cake, sandwich), recalled significantly more words than Ss given the list in a random order.

The expected first-order interaction between grade level and ordering condition was not significant, nor was that between grade level and superordinate condition. The three-way interaction of grade, sex, and ordering condition was significant, however; it is presented graphically in Figure 5-1. This figure shows that the differential advantage of ordering the lists for the older Ss held only for the girls. A test of the simple effects for boys and girls separately confirms this interpretation. For both boys and girls the effect of the ordering condition remains significant ($F = 7.75$, $df = 1,112$, $p < .01$; $F = 7.85$, $df = 1,112$, $p < .01$). For the girls, however, there is a significant interaction between grade level and ordering condition ($F = 9.12$, $df = 1,112$, $p < .01$), whereas for the boys there is not ($F = 1.31$, $df = 1,112$, n.s.). That is, for the boys the contribution of ordering the lists was essentially constant over the two grade levels.

Recall scores increased significantly over the four recall trials ($F = 42.0$, $df = 3,336$, $p < .001$). Moreover, this effect was more marked for the ordered lists, as shown by the significant interaction between ordering condition and number of trials ($F = 2.7$, $df = 3,336$, $p < .05$).

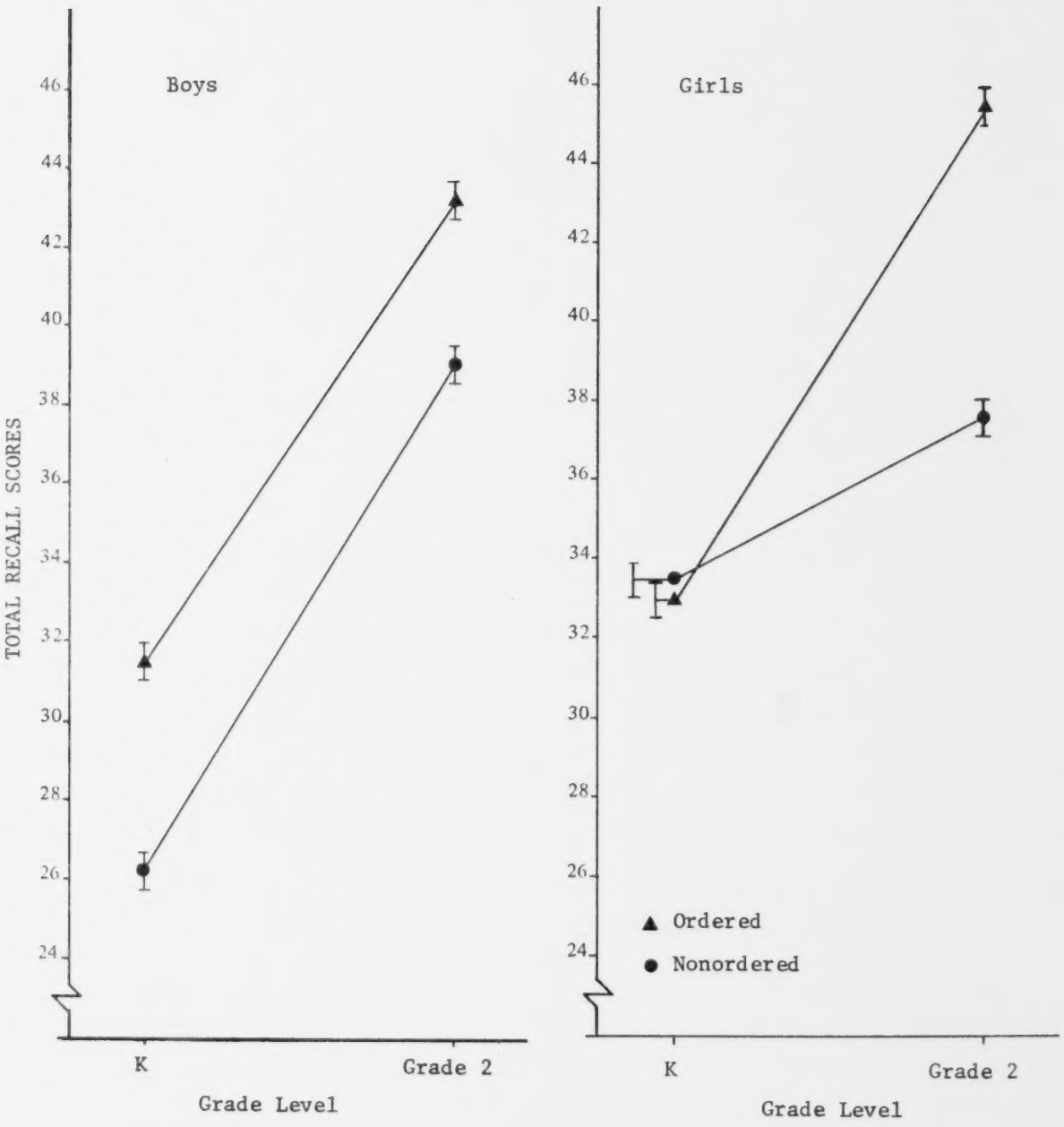

Fig. 5-1. The effects of ordering on recall as a function of sex and grade level.

This relationship is shown graphically in Figure 5-2. A trend analysis showed that the lines in this figure have significantly different slopes ($m_1 = .93$, $m_2 = .61$, $F = 4.56$, $df = 1,336$, $p < .05$).

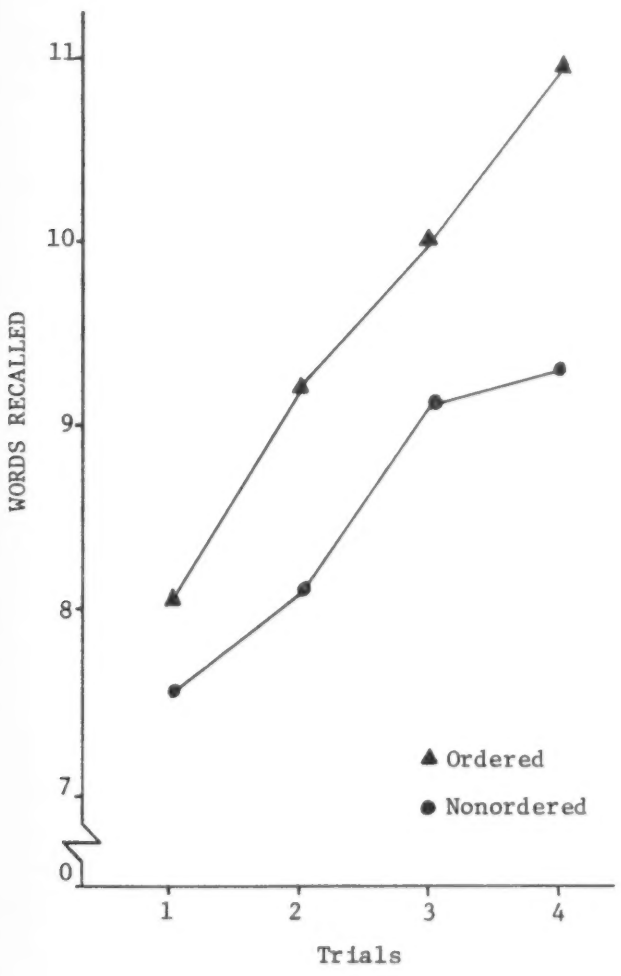

Fig. 5-2. The effects of ordering on recall as a function of number of trials.

The provision of the superordinate terms affected recall across trials, but only in interaction with grade level ($F = 4.0$, $df = 3,336$, $p < .01$) and sex ($F = 3.1$, $df = 3,336$, $p < .05$). Figure 5-3 shows that the provision of superordinate terms had a positive effect, primarily for the boys, but only in the later trials. This interpretation is confirmed by an analysis of the simple effects, which showed that the superordinate-by-trials interaction was significant for the second-grade children ($F = 7.92$, $df =$

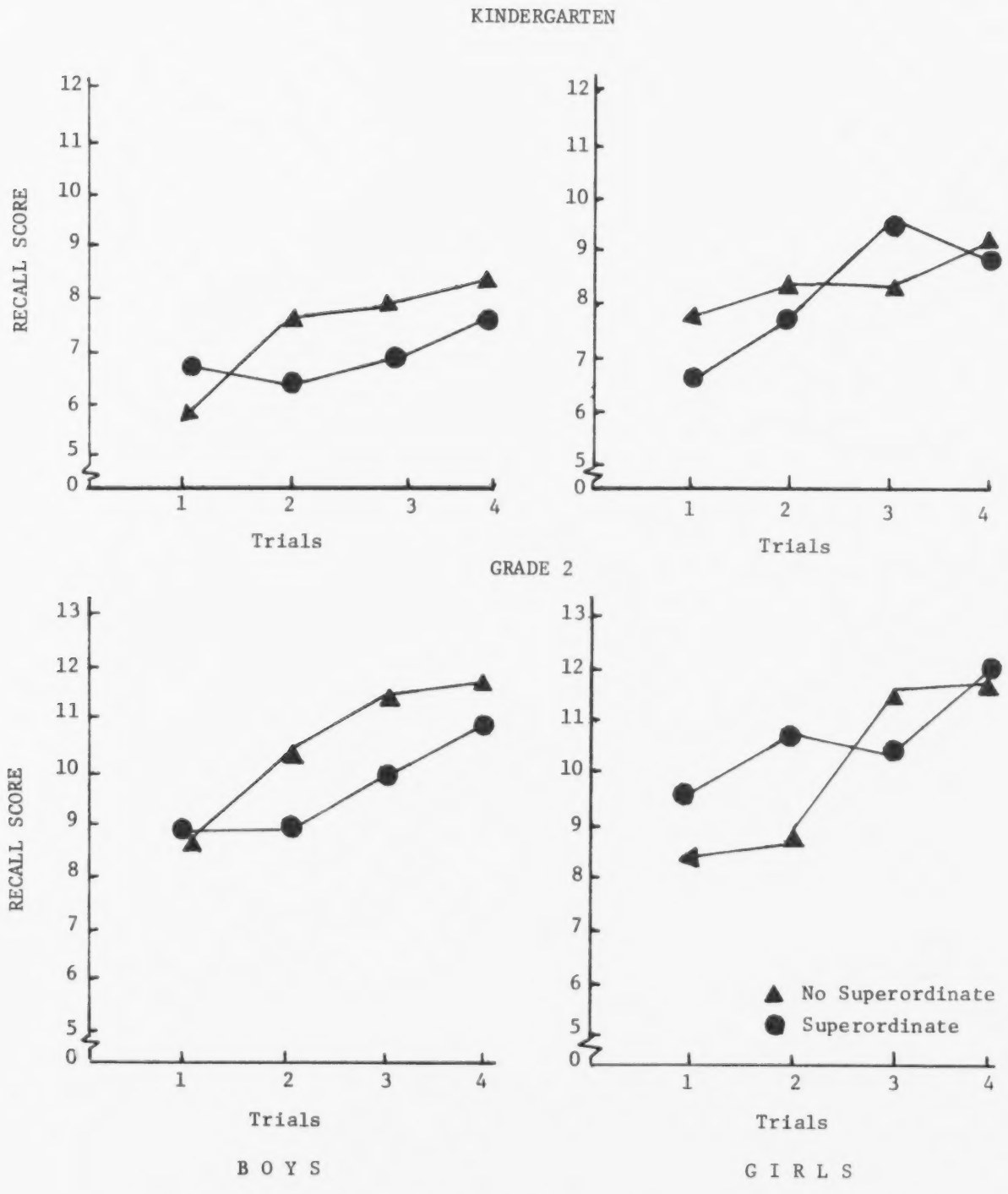

Fig. 5-3. The effects of superordinate terms on recall as a function of sex, grade level, and number of trials.

3,336, $p < .001$) but not for the kindergarten children ($F = 2.41$, df = 3,336, n.s.). Similarly, the superordinate-by-trials interaction was significant for the boys ($F = 7.75$, df = 3,336, $p < .01$) but not for the girls ($F = 1.28$, df = 3,336, n.s.).

Since the effects of the experimental variables were found to have increased over trials, the fourth-trial scores were examined to determine the differential effects of ordering and superordination as a function of grade and sex. As Figure 5-4 shows, for the second-grade boys, unlike the girls, the presence of the categorizing label compensated for the random

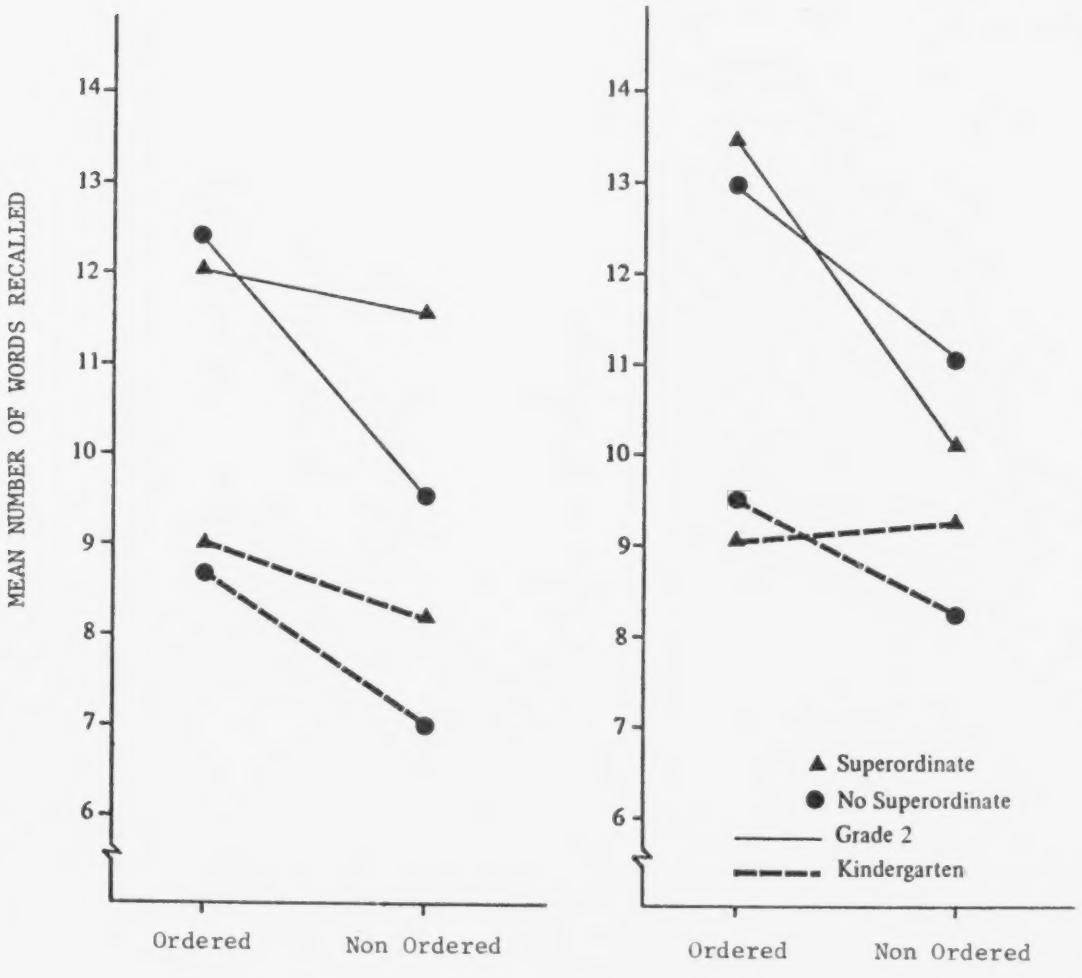

Fig. 5-4. Mean number of words recalled on Trial 4 as a function of sex, grade, ordering, and superordination.

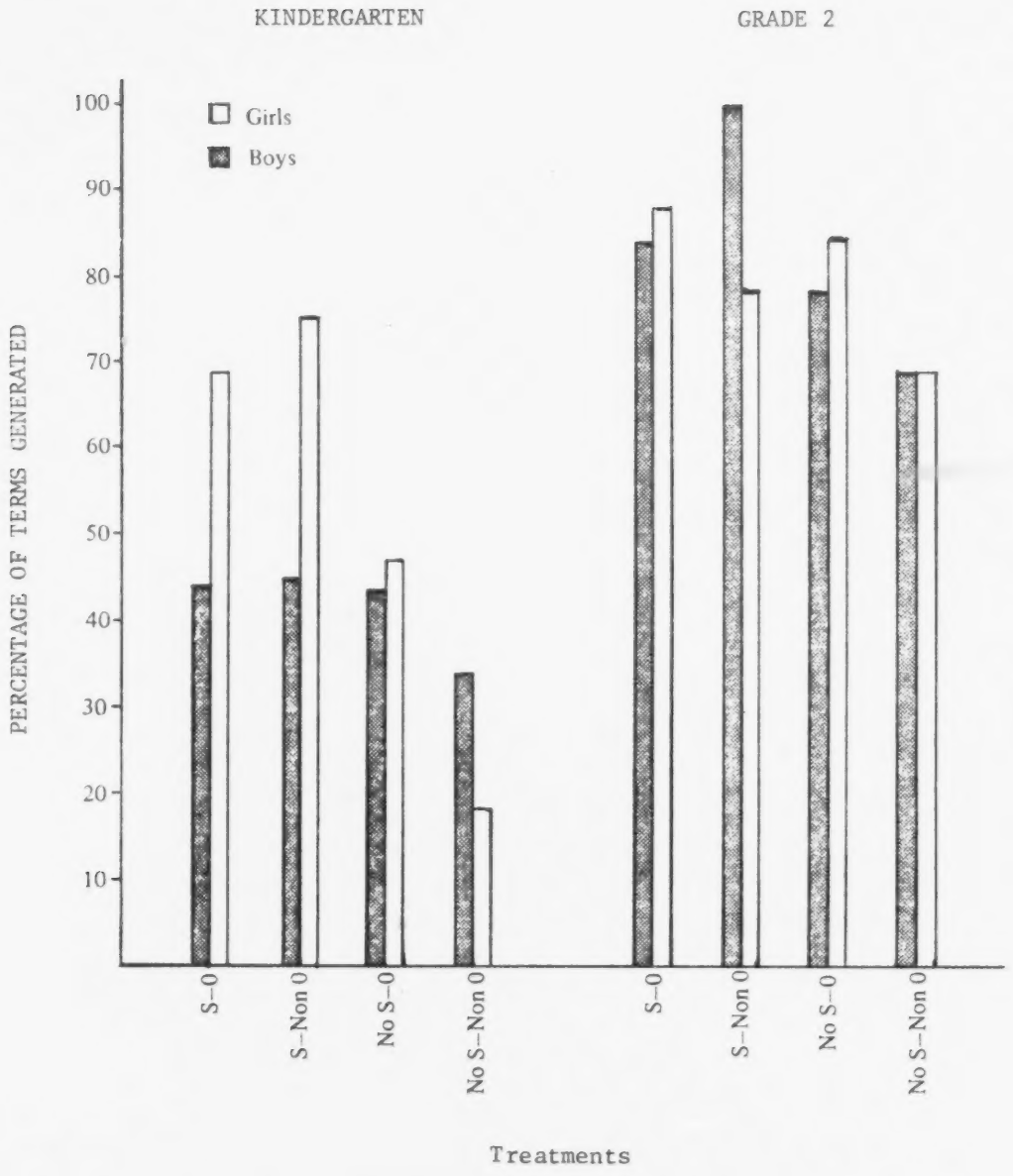

Fig. 5-5. Percentage of superordinate terms generated by S as a function of treatment condition, sex, and grade level.

ready provided by the ordering condition. That is, boys learned the list better with the help of either the ordering condition or the superordinate condition. For the older girls, the superordinate condition helped only in combination with the ordered condition.

It appears that the ordering condition had no effect on the performance of the younger girls. We do not have a treatment that was more appropriate for them, however. While a replication of the experiment is underway (Olson, Hildyard, & Crosbie, in preparation), the modest confirmation of the aptitude-treatment interaction provided by the girls in this experiment suggests that such an interaction may exist more generally, but that it failed to appear in the present experiment. Possibly the structure of the word list, which was expected to be useful to the older children, may have been circumvented by some other form of organization. Bower, Clark, Lesgold, and Winzenz (1969) have recently offered evidence that hierarchical organization of the sort involved in this experiment does substantially increase recall, but that simpler associative linkages are an effective alternative. In that experiment the first of these alternative forms of organization was based on the logical principle of class inclusion (e.g., metals--copper); in the second, it was based on simple associations between pairs of items (e.g., milk--cookie). On the basis of the theory previously discussed, only the former type of hierarchy would be expected to be developmental; the latter might be constant across various ages. If Ss in this experiment were organizing the word list on the basis of simple associative systems, a strong interaction between stage of development and form of treatment would not appear. The fact that it did appear for the girls suggests that at least some children must have used the class-inclusion system.

It follows, too, that a critical test of this interaction would have to differentiate materials that could be organized by association from those that could be organized by class-inclusion relationships. The interaction should appear in the latter case, but not in the former. An even simpler means of examining this interaction is suggested by the design of the preceding experiment. Stage of development could be expected to interact with the conceptual grouping condition of that experiment, but not with the narrative condition.

On the whole, however, it is obvious from both Experiment 4 and Experiment 5 that the effects due to the different recodings of the messages are consistently more important than the effects due to any of the summary indices of ability used in these studies. The empirical investigation of the factors that may be used to increase the amount of information children can comprehend and recall from various messages dealing with various types of information remains an important field of study. Bower (1970), for example, has provided dramatic evidence of the gains in learning that can result from the use of various mnemonic devices.

Although the evidence presented here indicates the relatively minor effects attributable to the pattern of aptitudes of the learners, including their stage of development, it must be obvious how closely each step in the instructional process must be tuned to the requirements of the

order of the words in the nonordered list. Thus, though the mean recall score was very low (9.5 words) when neither factor was present, recall was significantly higher in the ordered condition (12.4 words), and also in the superordinate condition (11.5 words). The combination of both factors resulted in a score (11.9 words) similar to those resulting from either factor alone.

For the second-grade girls the effect of ordering remained the most important (see Fig. 5-1). For them, furthermore, the provision of the superordinate did not compensate for the lack of ordering; it tended rather to make the situation worse. When superordinate terms were used with the ordered lists, the second-grade girls recalled significantly more words (13.6) than when these terms were used with the nonordered lists (10.1 words) ($t = 2.6$, $df = 14$, $p < .01$).

Some evidence of the developmental shift in the role of the superordinate terms is provided by the children's production of these terms at the end of the fourth recall trial. As Figure 5-5 shows, the second-grade children were much more successful in generating these terms--both in the superordinate condition and in the no-superordinate condition--than the younger children.

Discussion

First we shall consider the features of the treatments that were conducive to a higher level of recall. The significant effects of ordering the list in terms of conceptual categories in this experiment support the preceding experiment's finding that categorization is an important means of recoding information to increase recall. Moreover, since in Experiment 4 the effects of the categorization condition were not, statistically speaking, significantly different from the baseline condition, the support added by this data to the earlier conclusion is all the more important. As for the use of superordinate terms for these groupings to aid recall, it had a significant effect only for the boys, and then only in the later trials in the series. The general conditions under which the provision of such terms could be expected to be beneficial remain uncertain.

Next we shall consider the interaction between these treatment conditions and the stage of intellectual development of the learners. It was hypothesized that the ordered lists, with their intrinsic conceptual structure, would be useful to children who had mastered the concrete operation of class inclusion (second-grade children), but irrelevant to children who had not (kindergarten children). That theory held up very well for the girls: the ordering of the list had a substantial effect for the older girls and no effect for the younger girls, as Figure 5-1 shows. For the boys, however, there was no such distinction: the ordering of the lists was just as beneficial to the younger boys as to the older.

The effects of the superordinate terms are confusing. Later in the series of trials, at any rate, they appear to have provided the older boys with a means of structuring the word list, if that structure was not al-

listening or reading child. How a teacher should proceed cannot be determined by any simple pretest measure; rather he must continuously be thinking about what the child has already heard, what the child's conception of the task is, what questions the child may have in mind, and so on. The lessons must be tuned to the requirements of the child at each instant in the communication process, not to some objectively predetermined characteristic of the child. Whether the teacher says "yes" or "no" in any given instance, for example, must always depend on the question the child at that moment is considering.

EXPERIMENT 6: THE TEACHER'S ABILITY TO ENCODE MESSAGES

In Experiment 1 it was apparent that some teachers were much more successful at encoding messages than others. But that study offered no evidence about the consistency of this effect or its relation to other factors. Do such differences appear in more complex communication tasks? Is there any consistency from task to task? That is, are the teacher-encoders who are more successful at communicating one kind of information also the ones who are more successful at communicating other kinds? How is the ability to encode a successful message related to other skills or abilities?

This experiment, then, was an attempt to investigate further what is involved in the successful communication of information, that is, in instruction. To this end, teachers were asked to teach two kinds of tasks--a verbal task and a geometric task--to children at two different grade levels, grades 6 and 12. Again a communication-games format was used in which the teacher initially taught the lesson to one child, from whom he was separated by an opaque screen. The teachers' messages were tape-recorded and subsequently played to groups of children, whose performance on the tasks was then taken as an index of the effectiveness of the teaching. Finally, an attempt was made to account for differences in the effectiveness of different messages in terms of such variables as the length of the message; the teacher's characteristics, including amount of teaching experience, type of experience, sex, and ability to take another's viewpoint (as assessed by Hunt's Conceptual Level test); and the IQ of the students.

The decision to use a live listener during the recording of the lessons was based on the fact that in Experiment 1 some encoders seemed to have difficulty taking the task seriously and would inject comments to E and so on. Hence in this experiment "stooge" decoders of the same sex as the teacher were used during the encoding session.

Subjects

The teacher-encoders in this experiment were 35 teachers, 22 male and 13 female, enrolled in graduate courses in the Ontario Institute for Studies in Education summer session. The "stooge" decoders who actually listened to the teachers while the messages were being taped were a boy and a girl at the sixth-grade level and a boy and a girl at the twelfth-grade level. The actual decoders of the messages were 210 students, half from sixth-grade classes and half from twelfth-grade classes in suburban Toronto schools.

Materials

The verbal task (V) required the teacher to teach the following list of twelve singular, concrete nouns so that the student would be able to recall them an hour later: potato, gasoline, hill, milk, field, pill, forest, manure, hamburger, perfume, horse, rose. Clearly there are sev-

eral possible ways of structuring this material to make it easy to remember.

The geometric task (G) required the teacher to describe a series of four connected rectangles in such a way that the student, given one of them, could accurately draw the remaining three from the teacher's description. The model viewed by the teacher is shown in Figure 6-1; the one given to the student is shown in Figure 6-2.

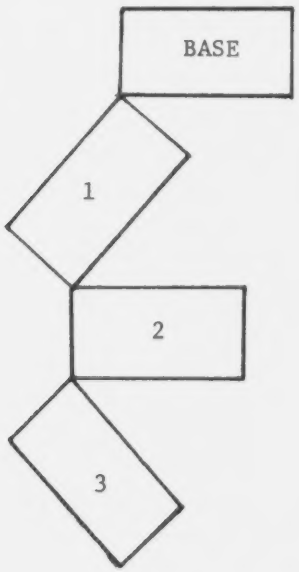

Fig. 6-1. Teacher's model for the geometric task.

Fig. 6-2. Model shown on student's answer sheet.

The conceptual-level test developed by Harvey, Hunt, and Schroder (1961) was administered to all encoders. This sentence-completion test, designed to measure a person's ability to take an abstract point of view, is composed of such items as: "When I am not sure . . . ." "When someone disagrees with me . . . ." In similar contexts Hunt (1971) found that Ss having a high conceptual level are more able to adapt flexibly to unfamiliar, unstructured situations.

Procedures

Encoding phase. There were two teaching tasks, each to be taught at two grade levels; hence each teacher taught four lessons. The order both of grade level and of task taught first by a given teacher was varied in four different ways to control for practice effects, as follows: 6G, 6V, 12G, 12V; 6V, 6G, 12V, 12G; 12G, 12V, 6G, 6V; 12V, 12G, 6V, 6G. The teachers were not told at the outset that they would have to teach the same lesson at two different grade levels. The experimental session proceeded as follows.

Each teacher-encoder was taken individually to a laboratory room, where he was seated at one end of the table facing a wooden dividing screen measuring two by three feet. He was told:

> We are trying to determine exactly how a teacher transmits information to a pupil. To study this, we have designed two separate tasks for you to teach to a boy (girl) in Grade 6 (12).
>
> The first (second) task is a verbal task in which you will teach this list of twelve words to the student. The words do not have to be taught in this particular order, so you may change it if you wish. The important thing is to teach the student so that he (she) will be able to remember these words one hour from now. He (she) will not be allowed to write down the words as you teach. Meaning and spelling are not important.
>
> The second (first) task is a geometrical task in which you will have to describe this drawing of rectangles [while showing him Figure 6-1] so that the student is able to reproduce them accurately. For this task, the student will have a sheet of paper with a rectangle at the top--like this [while showing him Figure 6-2]. As you are talking, he (she) will be drawing what you describe, on the page.
>
> All your instructions to the student are going to be tape-recorded. Would you please talk clearly and distinctly. The student will sit in this desk behind the barrier so you will not be able to see each other. Also, the student will be told not to talk to you or to answer questions. It should take you about fifteen minutes to teach these two tasks. Have you any questions?
>
> I will now give you a minute to plan how to teach these tasks while I go and bring in the student. I will tell him (her) only that there is a teacher in here who is going to teach him (her) a couple of tasks. He (she) is to listen carefully and do exactly what you tell him (her).

The "stooge" decoder was then brought in, introduced, seated on the opposite side of the dividing screen, and informed of his role as a listener. The teacher was then requested to begin teaching. Separate tape recorders

were used to tape the geometric and verbal messages. When the teacher had finished, he was told:

> Now I am going to ask you to teach the same tasks again, only this time to a boy (girl) in Grade 12 (Grade 6). Exactly the same procedure will be followed as before; that is, you will teach the verbal (geometric) task first and the geometric (verbal) task second.
>
> I will go and instruct this new student as I did the other one, and you may have one minute to plan what you will teach.

The second "stooge" decoder was then brought in and introduced, and the teaching recommenced. After the teachers had completed the four tasks they took the conceptual-level test and filled out a questionnaire requesting information on type and amount of teaching experience, guidance and counseling experience, and type of academic background--arts or sciences.

Decoding phase. The tape-recorded lessons from the 35 teachers were subsequently played to the real decoders. Three children served as decoders for each lesson; thus for each encoder there were six decoders, three at each grade level. The taped messages were played in the same order for all decoders: the verbal message first and the geometric message second. After the geometric message the children were asked to write down as many of the twelve words as they could remember from the verbal message. The specific procedures were as follows.

Groups of three Ss drawn at random from the list of participating children provided by the school were taken to the testing room, which contained three desks placed around the two tape recorders and about six feet away from them. E explained to the children that some experiments were being done to find out how teachers taught different things. The children were further told that the investigators had tape-recorded some teachers teaching and now wanted to find out what children learned from those lessons in order to find out whether some teachers were better than others. Just prior to playing the first lesson--the verbal one--E pointed out that the children would have to remember what they were taught; for the second lesson, the geometric one, E gave them the appropriate answer sheet (Figure 6-2) and asked them to draw what they were told to by the teacher. After the geometric task E gave the Ss an answer sheet with twelve blanks and asked them to write down as many of the original list of words as they could remember, in any order.

The IQ scores of all decoder Ss were obtained from school records.

Results and Discussion

Complete records including encoding transcripts, background data, and decoder performance were obtained for only 28 of the 35 original en-

coder Ss. These data were then analyzed in two ways--by analysis of variance and linear regression analysis.

Separate four-way analyses of variance were conducted first for the verbal task data and second for the geometric task data. In both cases the four factors were as follows: order (the tasks were taught in four different orders); encoder (there were seven encoders for each of the four orders, 28 in all); decoder (there were three decoders for each encoder); grade level (decoders were either sixth graders or twelfth graders). The results of this analysis for the verbal task data are presented in Table 6-1. This analysis shows that three of the factors ex-

TABLE 6-1

Four-Way Analysis of Variance for the Effects of Encoder, Decoder, Grade Level of Decoder, and Order of Tasks on the Decodability of Verbal Messages

Source	df	SS	MS	F
Order (O)	3	73.21	24.40	34.16**
Error	8	5.71	.71	
Grade (G)	1	174.05	174.05	95.13**
Encoder (E)	6	58.29	9.71	5.31**
Decoder (D)	2	3.76	1.88	
E x G	6	36.57	6.10	3.33**
E x O	18	120.33	6.69	3.65**
G x O	3	41.97	13.99	7.65**
D x O	6	1.95	0.32	
D x E	12	17.82	1.49	
D x G	2	1.86	0.92	
E x G x O	18	57.57	3.20	1.75*
D x O x E	36	57.13	1.59	
D x O x G	6	3.19	0.53	
D x G x E	12	36.89	3.07	
D x O x G x E	36	73.39	2.04	
Error	104	190.29	1.82	

*p < .05
**p < .01

amined--the order of the tasks, the encoder, and the grade level of the decoder--had a significant effect on the decoder's performance on the verbal task. That the decoder himself was not significant is probably due to the nature of the statistics: since there were only three decoders for each message, and since the three were different for each message, the differences between decoders tended to cancel each other out. A fairer estimate of the decoder effects will be obtained from the linear regression analysis.

Some orders of teaching the four lessons yielded higher decoder performance than others, as Figure 6-3 shows. Clearly, however, this effect is primarily due to the significant interaction between grade and order, as can also be seen from Figure 6-3. Thus for the twelfth-graders

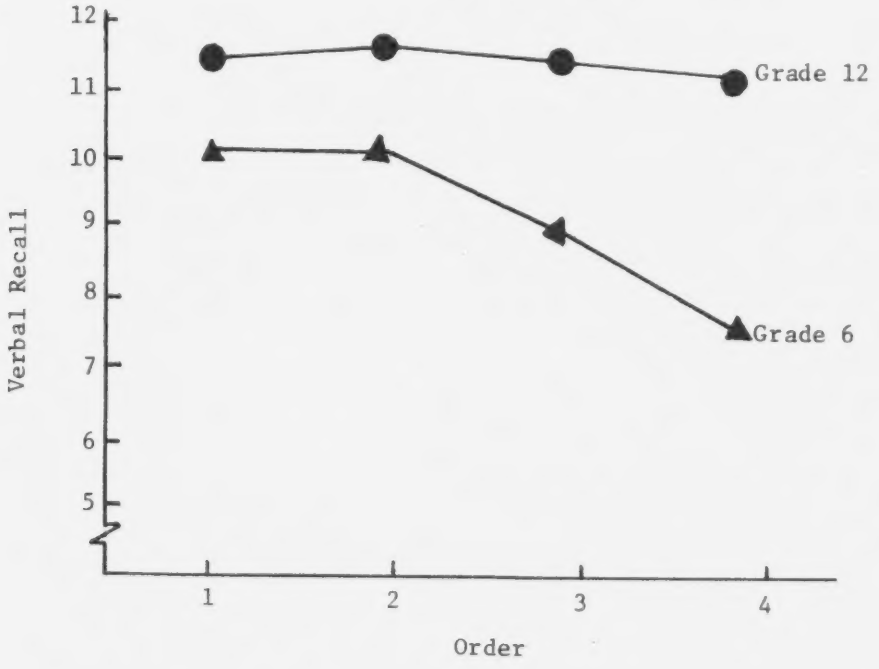

Fig. 6-3. Grade-order interaction on the verbal task. (Order 1: 6G, 6V, 12G, 12V. Order 2: 6V, 6G, 12V, 12G. Order 3: 12G, 12V, 6G, 6V. Order 4: 12V, 12G, 6V, 6G.)

the effect of order was very small, whereas for the sixth-graders poorer performance resulted from orders 3 and 4, in which the lesson was taught to the sixth grade after it had already been taught to the twelfth grade. The significant effects of the decoder's grade level are also clearly shown in Figure 6-3: the twelfth graders consistently outperformed the sixth graders. If it is assumed that the quality of the lessons did not vary

for the two groups, this difference can be interpreted as a decoder effect; that is, one can conclude that the decoder was a significant factor in the decodability of the messages.

The significant differences among the 28 different encoders may best be determined from the encoder-grade, encoder-order, and encoder-order-grade interactions, all of which are significant (see Table 6-1). The first of these interactions indicates that some encoders were more sucessful with one grade level of decoder than with the other. These patterns of interactions are much clearer in the linear regression analysis.

Almost identical effects were obtained for the performance on the geometric task. The results of the four-way analysis of variance for this task are presented in Table 6-2. Again there are significant effects for order

TABLE 6-2

Four-Way Analysis of Variance for the Effects of Encoder, Decoder, Grade Level of Decoder, and Order of Tasks on the Decodability of Geometric Messages

Source	df	SS	MS	F
Order (O)	3	111.29	37.10	4.79*
Error	8	61.90	7.74	
Grade (G)	1	971.52	971.52	51.95**
Encoder (E)	6	146.32	24.39	1.30
Decoder (D)	2	32.33	16.17	
E x G	6	379.73	63.29	3.38**
E x O	18	1084.30	60.24	3.22**
G x O	3	163.10	54.37	2.90*
D x O	6	29.57	4.93	
D x E	12	241.25	20.10	
D x G	2	5.19	2.60	
E x G x O	18	600.99	33.39	1.79*
D x O x E	36	674.85	18.75	
D x O x G	6	49.48	8.25	
D x G x E	12	191.06	15.92	
D x O x G x E	36	782.94	21.75	
Error	104	1944.76	18.70	

*p < .05
**p < .01

and for grade level of the decoder, and significant interactions between encoder and grade, encoder and order, and encoder, order, and grade, indicating the importance of encoder effects and a possible superiority of some encoders for some grade levels of decoders. The effects of grade and order as well as their interaction are shown in Figure 6-4.

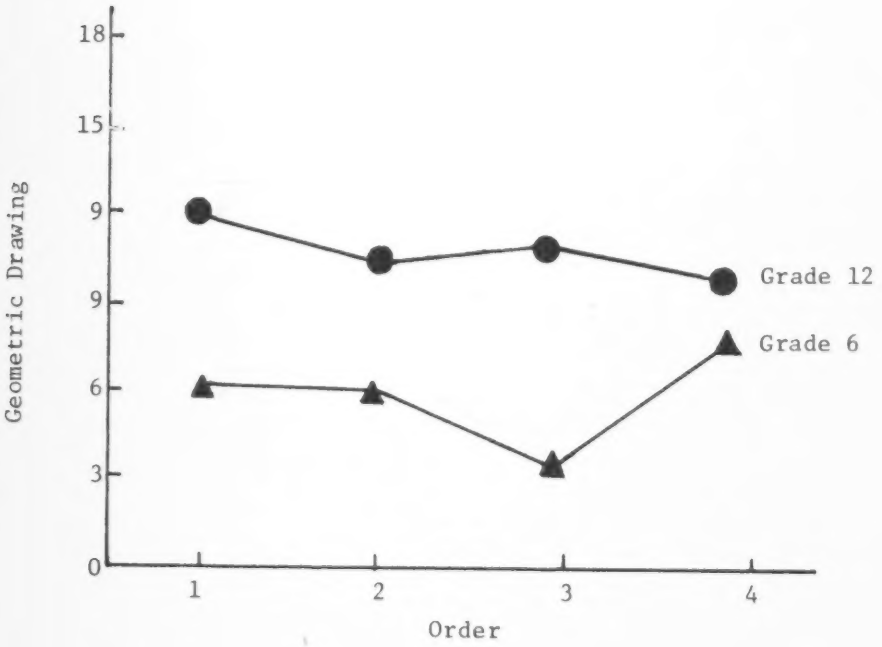

Fig. 6-4. Grade-order interaction on the geometric task. (For order key see Figure 6-3 above.)

Predictably, the twelfth graders again outperformed the sixth graders. Unlike the verbal task, the geometric task was apparently better taught the second time through: thus the twelfth graders performed better on this task when their teachers had already taught it to sixth graders, and vice versa. That explanation, however, fails to account for the poor performance of the sixth graders in order 3. If Figure 6-4 is compared with Figure 6-3, it is clear that for the sixth graders there is a marked cross-over for the verbal and geometric lessons in orders 3 and 4, the orders in which the sixth graders were taught after the twelfth graders. In both of these cases the first lesson taught to the sixth graders was not as successful as the second. It is reasonable to infer that teachers had some initial difficulty in adjusting their lessons to the requirements of the young decoders.

The results from the linear regression analysis shall be considered in two parts: first, the correlation among the various factors examined in the study, and second, an assessment of the variance in the decoder

performance that may be attributed to these factors. For the purposes of this analysis the performance of the three decoders of each message was averaged to yield one decodability score for each message. The IQs of the decoders were similarly averaged. A correlation matrix for the sixteen variables considered in this study is presented in Table 6-3. The first four variables are the decoder performance on the two types of tasks for the two grade levels; the next two are the IQs of the two grade levels of decoders; the next four are the lengths of the messages (by number of words) for the four lessons. The six remaining variables in the table pertain to the characteristics of the encoders: conceptual level, years of teaching experience, academic background (arts or sciences), sex (male or female), clinical or counseling experience (yes or no), and grade level taught (elementary or secondary). Since the last four variables are binary, each value (e.g., male or female) was treated as a separate factor in the analysis, yielding a total of twenty factors. For convenience these symmetrical values have been omitted from Table 6-3.

The statistical significance of the values in Table 6-3 can be assessed by the Fisher and Yates Table of Critical Values for df = 26; any value above 0.37 could be considered significant. This critical value is probably too high, however, in view of the fact that three decoder scores were used to arrive at each of the decoder scores used in the analysis. For example, with three times as many observations, a critical value of only 0.195 would be required for significance. For this reason we have taken any value above .31, the critical value for a one-tailed test, as significant.

Table 6-3 reveals some interesting patterns. The consistently positive intercorrelations among the first four factors indicate the consistency of the teachers' ability to generate an effective lesson. However, the positive correlations between sixth-grade verbal and sixth-grade geometric scores may be due either to the fact that the two lessons were taught by the same teacher or to the fact that they were decoded by the same children. To determine whether there was a consistent teacher effect, a partial correlation was calculated between 6V and 6G with the effects of the decoders' IQ partialled out, yielding a correlation of .25, which is not significant ($t = 1.27$, $df = 25$, n.s.). Similarly, the correlation between twelfth-grade verbal and geometric scores may be due either to the consistency of the teacher effect or to the fact that the lessons were decoded by the same students. Another partial correlation showed that with the effects of the twelfth graders' IQs partialled out, a significant correlation of .408 remained ($t = 2.24$, $df = 25$, $p < .05$).[1]

A further analysis shows that the effects due to the encoder may be even more substantial than the above analysis indicates. Whereas the above analysis is based on an average of the three decoder scores per message, in this analysis the set of correlations was recalculated, using individual scores on the verbal and geometric tasks and IQs for both grade

---

[1]The authors are indebted to Barak Rosenshine for this suggestion.

TABLE 6-3

Correlation Matrix for Sixteen Encoder, Message, and Decoder Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Decoder scores															
6V 1															
6G 2	.36														
12V 3	.53	.25													
12G 4	.31	.04	.44												
Decoder IQ															
6IQ 5	.30	.58	-	-											
12IQ 6	-	-	.20	.27	-										
Message length															
6V 7	.51	.45	.18	.37	.27	.43									
6G 8	.27	.20	.25	.29	.22	-.15	.30								
12V 9	.34	.48	.03	.04	.29	.39	.61	.32							
12G 10	.26	.12	.02	.11	-.10	.13	.42	.25	.41						
Encoder characteristics															
Conceptual level 11	.36	-.02	.09	.03	.16	-.03	.29	.27	.17	.11					
Years experience 12	.18	-.09	.10	-.06	.26	.05	.01	-.14	.00	-.05	.13				
Academic background 13	.14	.01	.29	.12	.10	-.07	.05	.02	-.12	-.09	.07	.02			
Sex 14	.05	.13	-.21	.07	.09	.25	.40	.33	.49	.04	-.06	-.08	-.33		
Clinical experience 15	.05	.14	.28	.31	-.13	-.11	-.05	.02	-.15	-.19	-.07	.09	-.04	-.12	
Grade level taught 16	.05	.01	-.12	-.20	-.03	-.06	-.06	-.04	.10	.06	.30	-.03	-.17	-.33	.10

levels. This set of correlations is presented in Table 6-4. The results parallel those in Table 6-3, except that the effect tends to be more significant, owing to the larger degree of freedom. The partial correlations to remove the effects of decoder IQ yield a significant correlation between

TABLE 6-4

Correlation Matrix for IQ and Individual Decoder Score on Verbal and Geometric Tasks

	Grade 6			Grade 12		
		1	2		1	2
Verbal	1			1		
Geometric	2	.32**		2	.36**	
IQ	3	.37**	.34**	3	.19	.32**
	**df = 95, p < .01			**df = 82, p < .01		

verbal and geometric performance both for the sixth graders ($r = .22$, $t = 2.24$, $df = 95$, $p < .05$) and for the twelfth graders ($r = .32$, $t = 3.22$, $df = 82$, $p < .01$). Thus there is some consistency for the decoders across tasks that cannot be attributed to their IQ. Obviously the consistency of the encoders is a likely explanation.

Returning to Table 6-3, it may be seen that there is a significant correlation of .53 between the scores of the sixth graders and those of the twelfth graders on the verbal task. Since the only thing common to these groups is their teachers, this correlation may be attributed with some certainty to the teachers' consistent encoding ability on repeated attempts on the same task; the correlation of .31 between sixth graders' scores on the verbal task and twelfth graders' scores on the geometric task suggests a consistent encoding ability across different teaching tasks.

The consistently positive correlations among the lengths of the four messages also indicate an encoding consistency across tasks, aside from the criterion of decodability. That is, encoders who generated a long message for one task and grade level also tended to generate long messages for other tasks and grade levels. Thus there does appear to be some teacher consistency across lessons and tasks, at least for the materials used in this experiment.

The IQs of the decoders appear to be consistently related to their performance, but these values reach significance only for the sixth graders on the geometric task.

Message length is consistently positively related to decoder performance. It was an important factor in the decodability of the verbal messages for the sixth graders, yielding a correlation of .51; but the correlation drops to .03 for twelfth graders on those tasks. That is, longer messages were an asset to the younger decoders but not to the older ones. Furthermore, the mean length of the verbal messages for the sixth graders (210.36) was significantly greater than for the twelfth graders (164.71) ($T = 2.08$, $df = 33$, $p < .05$).

The conceptual level of the encoders is significantly related to decoder performance on the sixth-grade verbal task, and it approaches a significant relationship with the lengths of the messages to the sixth graders. The clinical experience of the encoders is consistently positively related to decoder performance, but the relationship is significant only for the twelfth graders on the geometric task. Although the sex of the encoders is not significantly related to the decodability of their messages, it is significantly related to the length of both the verbal and the geometric message for sixth graders. On the whole, the male encoders generated longer messages, if not more successful ones.

Four linear regression analyses were conducted, one for each lesson: sixth-grade verbal, sixth-grade geometric, twlefth-grade verbal, and twelfth-grade geometric. The analyses examined the extent to which the various independent variables--the length of the lesson, the mean IQ of the decoders, and the six characteristics of the teachers--were related to the decoders' performance. Of the six teacher characteristics, only clinical experience had a significant effect on decoder performance, and then only for the twelfth-graders (geometric task, $F = 4.94$, $df = 1/19$, $p < .04$; verbal task, $F = 3.22$, $df = 1/19$, $p < .08$). Hence, though there was some consistency in the way the tasks were taught, it is not explained by the variables considered in this study.

The linear regression analysis confirmed the relationship between message length and decoder performance on the sixth-grade verbal task ($F = 6.12$, $df = 1/19$, $p < .02$) as well as that between decoder IQ and performance on the sixth-grade geometric task ($F = 11.11$, $df = 1/19$, $p < .003$).

In an effort to determine the major differences between the best and the poorest messages, we examined the transcripts of the three best and the three worst sixth-grade messages as measured by decoder performance. Although no firm conclusions can be drawn from such a small sample, certain trends did emerge. Consider the following examples of successful and unsuccessful verbal and geometric messages. The first is a successful sixth-grade verbal message (mean decoder score = 11.7):

> Now, Eric, I have a list of words here. I'm sure you know what they mean. There are twelve of them and we're going to talk about these words for a few minutes. I'll repeat them to you, and then what I want you to do is be able to remember these words, not so much the meaning or the spelling but just be able to remember them and see if you can say them off by heart a little later. Now, first I'll go through the words, and then I'm

going to go through them again in a different order, and I'll try to group them so that they're words that go together meaning-wise. I'll just go through them first: potato, gasoline, hill, milk, field, pill, forest, manure, hamburger, perfume, horse, and rose. Now here's our twelve words. Now it's getting near dinner time. Maybe we'll take the foods out of our list first; it might help to remember them. Our first one was potato. Most people eat potatoes. We get French fries from them, so potato and hamburger go together. With it we might drink milk; that's all we have. So potato, hamburger, milk. That takes care of a quarter of them right there--three of them out of twelve. Now, if you happen to eat too quickly, or too much, you might have to take a pill, so maybe we'll just slip pill in there, since it's something we would have to eat in case we ate too many potatoes, drank too much milk or ate too much hamburger--we'd have to take a pill. That's four, or one-third. And we might go to a drive-in to eat this type of food, and if we went in a car, we'd have to use gasoline. So to get to the potato, hamburger, milk, and hopefully not the pill, we need gasoline. Now it might be out in the country. On the way we might see a horse, and that's another one of our words, and we have field and forest and hill. Now let's see--to get to our hamburger and potato and milk and pill, we use gasoline. Those are all our words, and on the way we may have seen a horse on a hill or in a field or in the forest, and that leaves three words. They all pertain to smell, two of them probably pleasant, the other one maybe not so pleasant. Rose, perfume, and manure. I think you'll find that you'll remember these words if you think of our little story about going to eat. We could probably fit the manure in with the horse, the rose in with the field, and they're both going to have something to do with perfume. One not so pleasant, and the other more pleasant. So to get to our meal we had to drive using gasoline--that's one word. Let's go over them. We used gasoline so that we could go and eat our hamburger and drink our milk, eat our potatoes. Hopefully we didn't eat too much so that we didn't have to take a pill, which is our fifth word. On the way we saw a horse on a hill, or maybe in a field, or maybe in the forest. We may have seen some roses or manure near the horse, and both of these might remind us of perfume. Now, as I say each word, you can think what part of the story it fits into. I'll read just the way I did before, but this time they should mean a bit more, particularly if you remember that instead of my just telling you twelve words, they're all part of the story, and if you remember the story you can remember the words. Potato, gasoline, hill, milk, field, pill, forest, manure, hamburger, perfume, horse, rose. Now see if you can tell yourself the story. Print each word that you think is one of the twelve in your story, one of the twelve I gave you, and see if you come out with the right words. I'll give you a minute to do that. He can do anything but write, eh, rather than talk to me? All right, close your eyes, then you'll have a good imagination, Eric, and you don't have to write; it's much easier to think about them. So go over the story. See

if you can get the twelve words. All right, I'll go over it once more, just so that we've got the right twelve words, because there are a few others in there. All right, we were feeling hungry so we drove out to get something to eat at the drive-in. We used our first word, gasoline. When we got there we ordered and ate potato, our second word; hamburger, our third; drank some milk, that's our fourth word. So we've got gasoline, potato, hamburger, milk--if we didn't eat too much we won't have to take a pill, that's our fifth word. Now on the way there or back, since it was out in the country, we saw a horse and a hill, saw a field, or we may have seen him at the edge of the forest. So we've got hill, field, forest and horse. Now near the horse we may have seen some manure, which has a perfume, or a rose, which also has some perfume. Now we have our last three words, manure, rose, and perfume. Okay?

Next is an unsuccessful sixth-grade verbal message (mean decoder score = 3.3):

Eric, I have a list of words that I want you to remember. I want you to think first of all of a pill. It's a very special pill because inside of this pill three types of foods are contained, a potato, milk, and a humburger. So the first group of words that I want you to remember are the words pill, potato, milk, and humburger. So if you just remember the word or the idea of a pill and inside of this pill three things: potato, milk, and hamburger, I think you'll remember them. The second set of words that I want you to remember is under the word or heading of smell. Under this smell or the smells are perfume, a rose, like the smell of a rose, field, sort of the smell you can smell from the field, forest, gasoline. So if you remember the words perfume, and the smell such as rose, field, forest, gasoline, you can remember them. So the words I want you to remember are perfume, rose, field, forest, gasoline. So the last group of words that I want you to remember are very easy because there's a pattern here. They are horse, manure, and hill. So you have three groups of words that I want you to remember. First of all, pill and under this heading of pill, potato-milk-hamburger. The next set is perfume, smells, and I want you to remember the word perfume, rose, field, forest, gasoline, and the last group of words, there's only three actually, horse, manure, and hill.

Next is a successful sixth-grade geometric message (mean decoder score = 14.0):

All right, Eric, you have your paper in front of you, you'll see a figure called a rectangle at the top. Now we're going to draw three more rectangles the same size going down from this one. Now I'll give you the instructions for each one, starting with the first one. Now each one is attached by a corner and they're on different angles--look something like a chain. The first one,

as you'll notice, is horizontal. In other words, it's parallel to the top or bottom of the paper. Now the first rectangle we'll start--don't draw yet--the first rectangle attached to the lower left corner--you got that? That's where we'll start. Now it's hanging in the lower left corner. I don't know if I can bring degrees into this or not, but presume he doesn't have the protractor for 45°. I see. Okay. All right, I want you to draw a line from that lower left corner of your rectangle and the length will be the same as one of the shorter sides of the rectangle, the same as the shorter side, and pointing down towards your lower right corner of the piece of paper. Now draw that one line first, same length as the short side of your rectangle, starting from your lower left corner and pointing down toward the lower right corner of the piece of paper. Now you have that line, that's going to be the top end of your new rectangle. Now I want you to draw the two longer sides down from that. Now you'll notice the rectangle has right angles, so you know which way to draw those. Same length as the rectangle that was printed on your page and going down from the line you first drew. Now you'll notice they're roughly pointing to the left lower corner of your paper; not quite, but they're over on the left side, but like railway tracks they're parallel, or they're the same distance apart; they don't come together or they don't go apart. Now all you have to do to complete your rectangle is draw the fourth side, and of course, it's the same length as the first line you drew. So you've drawn four lines, two short and two long ones, and that's your number one rectangle. So it's hanging from the lower left corner of the printed rectangle; it's going down to the left. Now your second rectangle is attached to the lower corner of the one you just drew. Now I want you to draw a line which will be the short side of the rectangle and just going straight down, and draw it from the lower corner of your rectangle. So that's the short side--it's a line the same length as the short side of the rectangle you drew, it's attached to the lower corner, and it goes straight down. Now I want you to draw your rectangle or attach the other three sides to that so that it goes out to the right, straight out to the right. Don't forget the two lines of the same length; they're like railway tracks of the same distance apart or parallel, and it goes straight out to the right, something like a flag. You put your fourth line or the short line to enclose your rectangle and it's vertical or up and down--it looks like a flag. Now the third rectangle--and all these are the same size, same shape--the third rectangle is attached to the lower left corner of the rectangle you've just finished and we'll draw the short side again first, but this time the line is pointing down towards the left corner. Now I want you to attach the two long sides, or two long lines to the short side that you just drew so that they're pointing down to the right, to the right and down. I shouldn't have to tell you which direction. I should just tell you right, because if they're the same as the first rectangle, they're right angles or a square angle. Now once you've got those two drawn, you just enclose the rectangle with the fourth line which is your

short side. Now that last one again is hooked onto the lower left corner of your second rectangle that you drew. You drew a short line going down and to the lower left corner of your paper; then you attached the two long sides so that they were pointing or drawn down and to the right from that line; and then the fourth side to enclose it.

The last is an unsuccessful sixth-grade geometric message (mean decoder score = 1.3):

All right. Susan, I'd like you to place the top and the bottom of the sheet so that the rectangle base is up at the top of the sheet. Now I want you to turn the sheet so that the broad side of the paper is down at the bottom of your desk. All right, the rectangle that is on your paper should be at your left-hand side and on the left-hand side of the paper. Now we're going to make three rectangles look like a W. You're going to take your pencil and place it on the bottom right-hand corner of the rectangle that's on the paper. Now you're going to draw a straight line running from that point about an inch and a half on a straight line like the left arm of the letter W. That is a long line. Now from the bottom of that line that you have just drawn I'd like you to draw a short arm the same length as the short arm of the rectangle that is already on your paper, and it is to be the short arm of the letter W. Now from the top corner of that line you've drawn I want you to draw a long line parallel to the first long line you drew. Now close the rectangle. All right. Now from the bottom line of the rectangle that you have just drawn, the line that is facing you that's closest to you, I want you to point, put your pencil on that corner and draw a short baseline to the right. It should be the same length as the short base line of the rectangle that's on the paper. Now, I want you to draw a straight line going towards the top of the paper parallel to the right-hand side of your sheet. It should be the same length as the long line on the rectangle that is already on the paper. From the point at the end of that line, draw another short line parallel to the first line of the second rectangle you're drawing. All right. Now draw a second long line parallel to the long line that you drew a few seconds ago. This should close off the second rectangle. Okay. Now, start at the bottom right-hand corner of the second rectangle, the rectangle you have just completed. Draw a line parallel, pardon me, not parallel, going off to the right-hand corner--top right-hand corner of your sheet, and it should go in the opposite direction to the first rectangle you've drawn. It should be the same length as the first rectangle on the paper. All right. Now, from the top point of that line I'd like you to draw a short line the same length as the first short line of your original rectangle that's on the paper going towards the bottom right-hand corner of the paper. All right. Now, I want you to draw a line from the bottom. From that line you've just drawn you're to draw another long line parallel to the first long line you just drew of this third rectangle, and it should be

> going from the right-hand side of the page down to the bottom of the page facing you. Now, close the rectangle. The point of the closing line of the rectangle should join the bottom right-hand corner of the second rectangle you've drawn. All right.

For both types of tasks, good encoders were markedly more explicit in setting out the objectives of the task and the general strategy they would use for teaching it. In other studies we have called this procedure "prestructuring." Other successful techniques are more specific to the type of lesson being taught.

For the verbal task, the good messages were generally longer than the poor ones, and they tended to be more repetitious. In the three best messages each critical word was repeated an average of 11.6 times; in the three poorest ones, only 4.5 times. Thus good teachers demonstrate Ebbinghaus's laws of list learning.

Equally as important are the effects of organization on memory. Good encoders restructured the verbal material in some manner, such as categorization, and then consistently used this recoding. One of the best encoders repeated each of the critical words a median of 23.5 times; yet the sucessful encoder whose message was transcribed above achieved a slightly higher level of recall with only 6.5 repetitions of each word by coupling repetition with an appropriate category system for grouping the words. Among the grouping procedures used by the encoders in this experiment were the following:

1. Categorization: "Now it's getting near dinner time. Maybe we'll take the foods out of our list first; it might help to remember them. . . . So potato, hamburger, milk."
2. Mnemonics: "Two words start with 'F,' and the two words are field and forest, field and forest; two words start with 'F.'"
3. Sentence grouping: "Well, I suppose maybe we could say that horse manure is found in the field, and I suppose it would make the rose grow if you put it on the rosebush, right? Horse manure, you find that in the field; horse is one word, manure is another word, in the field is good for the rose. Maybe we could make a sentence like that."
4. Narrative: "I think you'll find that you'll remember these words if you think of our little story about going to eat."

Good encoders generally used one or more of these recoding strategies, and used them in a consistent manner. Poor encoders generally used incomplete inconsistent recoding and were frequently ambiguous about which of the words in the longer sentences were the ones to be remembered.

The effect of recoding or conceptual grouping on recall is an important new topic in studies of verbal learning and recall. Mandler (1967) has shown that adults given a list of words to recall can usually recall about five. If they recode or categorize the words, however, they can recall about five words from each category, up to about five categories. He

concludes: "Memory is hierarchically organized into succeeding levels of categories and . . . list recall is a linear function of the number of categories used by the subject in organizing the list [p. 42]." Tulving (1962) has found that even in the free recall of unrelated words, adults subjectively organize the words, thus facilitating recall. The teachers in the present study apparently knew intuitively the effects of organization on recall, for many used this technique in generating their lesson.

In addition to these recoding strategies, many teachers concluded their lesson with a brief review, and a few asked the student to test himself on the words to be sure he knew what they were.

As for the geometric task, the standard deviation of decoder performance is substantially larger than for the verbal task at both the sixth-grade level (SD_G = 3.96, SD_V = 2.02) and the twelfth-grade level (SD_G = 3.59, SD_V = 0.67). Yet pinpointing the differences between successful and unsuccessful geometric messages is more difficult. To be sure, if one tries to draw the geometric patterns described in the messages reproduced above, it becomes obvious that the instructions in the successful message are much easier to follow than those in the unsuccessful message. But such obvious factors as length and the use of repetition seem irrelevant; instead, coherence and sequencing appear to be crucial. For example, many good encoders assigned numbers to the rectangles and signaled to the listener clearly when they completed one rectangle and began another. Poor encoders often mixed referents and failed to specify what they were talking about--for example, whether they were describing a particular line of a rectangle or the position of the rectangle as a whole. Others seemed to have had little feeling for the complexity of extracting spatial information from a verbal message, as in the following:

> Eric, I think you can estimate 2 3/4 inches; try anyhow. Usually from the tip of your thumbnail to the end of that knuckle if you put it down on that desk, is about an inch. Well, you put about 2 3/4 of those below the bottom left-hand corner of the rectangle that you already have.

Thus there remains the important and unsolved problem of determining the properties that make a message of this type easy to comprehend. Research in the related fields of readability (Taylor, 1953), psycholinguistics (Clark, 1969), and the processing of connected discourse (Frase, 1969), sidered above, may offer some help with this problem in the future.

Conclusions

The results of this study have added considerable weight to our earlier conclusion that there are marked differences among encoders in the way they generate a message or lesson, and that these differences have an important effect on the decodability or effectiveness of a lesson. This conclusion has been elaborated in four areas: teacher-encoder consistency, decoder effects and decoder-message interaction, encoder-decoder interactions or adaptability, and the predictability of teacher effects.

Teacher-encoder consistency. This study has provided evidence that there is some consistency in how a given teacher-encoder proceeds in teaching a given lesson to different audiences and even different lessons to different audiences. This consistency appears in the length of the message--the one feature of the message itself that was examined here--and more important, in the effectiveness of the message. Rosenshine (1969) recently reviewed the few studies that have been done on the consistency of teacher effects in ordinary classrooms. He concluded that though there are some consistent positive findings in short-term studies in which teachers taught the same material to different groups of pupils, there is no evidence for consistency across different lessons, even within the same subject area (cf. Fortune, 1967). The present study did yield such evidence, but in a severely restricted laboratory situation with little teacher planning time, brief tasks, and no feedback from the listeners.

Decoder effects and decoder-message interaction. The study of teacher effects is greatly complicated by their interdependency with those effects attributable to the child. This decoder effect is clearly shown by the consistently superior performance of the twelfth graders over the sixth graders when the two grade levels were given the same tasks. Furthermore, the correlation between the decoder's performance and his IQ, when calculated on an individual basis, was consistently positive and significant.

Encoder-decoder interactions. But the decoder effect is perhaps more substantial than any simple individual difference measure would suggest. When children know the requirements of the task they may compensate for failings in a teacher's message in unexpected ways. In the present experiment, for example, one teacher simply recited the list of words twice in a slow, measured pace, and the students compensated by rehearsing the words to themselves.

There are other lines of related evidence in the present study. One characteristic of the verbal lessons that helped younger decoders was the length of the lesson ($r = .51$), presumably because of the amount of repetition afforded by a long lesson. Yet for older decoders that relationship dropped almost to zero ($r = .03$). Thus a characteristic that makes a message appropriate for one audience does not necessarily make it appropriate for another. Our teachers appeared to know this fact, for they made the verbal messages for sixth graders significantly longer than those for twelfth graders. But if repetition or elaboration made the verbal lessons more effective, at least for the younger decoders, it appears to have been less relevant to the geometric lessons, for the correlations between message length and decoder success on the geometric task tended to be small and nonsignificant. Thus a strategy that is helpful for one type of lesson may be irrelevant for another type. One can conclude that message effects may depend both on the type of information to be transmitted and on the characteristics of the learner.

This interaction was described and examined in Experiments 4 and 5 in terms of aptitude-treatment interaction (ATI)--the hypothesis that some types of treatments may be more effective with certain types of children. The differential effects of length of message on sixth graders and twelfth

graders are indicative of such an interaction, but not conclusive, since the two grade levels heard different lessons. It would be interesting to play the sixth-grade messages to twelfth graders to see whether the correlation between length of message and decoder score did in fact fall to zero.

The interaction between encoder, lesson, and decoder may be even more subtle than has been indicated thus far. Not only must a message be generally appropriate to an audience having certain stable personal characteristics, it must convey the information appropriate to the requirements of the listeners for the task at hand. The appropriateness of a message must be reassessed at each point in the communication process. This continuing assessment is presumably made by a speaker according to feedback received from the listener. Thus the speaker-encoder continuously depends on the feedback from the listener, or some estimates of what that feedback is likely to be, to guide each step or phase in the communication process. Instruction, being a process of communication, must therefore incorporate a feedback loop.

This cybernetic model, which implies a complete interdependence of the speaker and listener, is well known in information theory (Wiener, 1954) and communication theory (Berlo, 1960). The important implication of this model is that there is no such thing as an ideal message as such: a message can be considered ideal only in relation to its hearers and their informational requirements for the tasks at hand. As Quine (1953) has pointed out, words have no absolute definition but are defined as required to facilitate communication in some context. But this complex point need not be belabored here. In the present study, the interdependence of speaker and listener appears in one piece of evidence: teacher-encoders used significantly longer and presumably more repetitious messages in the verbal task for younger decoders than for older decoders. That is, their knowledge of the decoders' age determined to some extent how they would proceed, even in the absence of continuous feedback. Furthermore, the teachers' adjustment to the younger listeners was appropriate, since the longer messages apparently facilitated the decoding process in their case. One aspect of good encoding is therefore the ability to size up the listener (Hunt, 1971). One can assume that if the teachers in our experiment had had access to continuing feedback from the listeners, they would have modified their lessons considerably more than they did.

Although the entire group of teacher-encoders tended to lengthen their lessons for the younger decoders, it could be expected that some would make a greater adjustment than others, owing to individual differences among teachers. It appears that this was indeed the case, in that the lengths of the various lessons taught by a given teacher were consistently positively correlated.

Predictability of teacher effects. We have indicated some differences among teacher-encoders and some evidence that teacher effects are consistent. This study also attempted to make a preliminary assessment of the

predictability of these effects from a set of independent variables measured at the time of the encoding sessions. None of these variables were consistently successful in this regard, but the clinical experience of the teachers did predict the effectiveness of their messages for twelfth graders. This finding may indicate that the teachers with greater clinical experience had the ability to make better assumptions about their listeners, "to take the role of the listener" (Flavell et al., 1968), in the absence of feedback. That this factor did not account for the effectiveness of the messages for sixth graders may be related to the fact that most clinical and counseling activities are carried out at the high-school level. It should be noted, too, that clinical background did not have its effect simply by increasing the message length, for there are negative correlations between these factors at the twelfth-grade level.

A second factor that may be relevant, though it was not significant in the linear regression analysis, was the conceptual level (CL) of the teachers. This factor correlated significantly with the performance of the sixth graders on the verbal task, and it was consistently positively related to message length. Since long messages were beneficial, especially for the sixth graders, it is possible that high CL teachers better anticipated this requirement of their listeners. Although this study did not examine the possibility, it is reasonable to suppose that high CL encoders would also be more capable of inventing the various recodings found in the analysis of the transcripts. If this were the case, it would follow that the high CL Ss were primarily responsible for the fact that verbal messages tended to be longer for the sixth graders. As mentioned previously, the effect of CL is independent of clinical experience and seems to act in a different way.

On the whole, the results for this area of the experiment are disappointing. Some factors in the experiment may have reduced the effects of the independent variables in question. First, the word list was easy, the mean score being 10.5 out of 12 words; thus there was little variance to be accounted for. Second, the geometric task was quite difficult, especially for the sixth graders, one-quarter of whom scored zero on this task. Third, the size of the sample was much smaller (N = 28) than is desirable for this type of analysis.

At the same time, it may be ill-advised to seek a set of independent variables that predict encoding ability, for three reasons. First, this procedure is actually a regression to the approach to teacher effectiveness that we abandoned at the outset of these studies--namely, that teaching ability can be accounted for by a set of ability or personality factors relating to the teacher. Second, the ability to generate an informationally adequate message appropriate to the audience involves a set of skills that cannot be reduced to combinations of traits. That is, the ability to generate effective messages must be attacked directly. Once the process is somewhat better understood, it may be possible to invent a measuring device that would assess these communication skills well enough to predict a teacher's effectiveness across a certain range of tasks. Third, an effective encoding always makes some assumptions about the decoders. As we have seen, some features of lessons that make them effective for one

group of decoders may be irrelevant to the needs of another group. Message length was such a factor in this experiment. Furthermore, encoding effectiveness frequently cannot be predicted at the outset, depending as it does on the encoder's continuous interaction with the decoder.

Teacher strategies. The range of recoding strategies devised by the teachers, given only one minute of planning time, was impressive. The strategies they used--conceptual grouping, narrative, mnemonics, repetition--are almost identical to the strategies proposed by the experimenter in designing Experiment 4 on the basis of previous research on factors influencing recall. That is, teachers appeared to know intuitively what psychologists have taken years to discover and demonstrate empirically. No attempt was made in this study to see how many of these recodings each teacher used; any one could have been effective for this task. It would be interesting, however, to find out whether some teachers can generate a large number of recodings for a given message in a shorter time than other teachers. Such a skill may be related to scores on some existing test, such as the conceptual-level test; alternatively, it would be easy to create a test to measure that skill directly.

One limitation in the performance of the teachers was that almost all of them accepted total responsibility for the child's learning, doubtless because they conceived of teaching in the linear or cause-effect sense that we abandonded in the preceding discussion. That is, the teachers assumed that what they did was the direct cause of the child's performance; the child's role was to follow passively the teacher's initiatives. In this experiment a teacher might have said (but none ever did), "I have a list of words here that you are to learn. I will read them to you slowly and you say them over to yourself three times after I say each word." Or, "As I read each pair of words, I want you to try to think of a sentence that will connect them." Or, "Try to make up a story using these words; your story will help you recall them at the end of the experiment." Bower and Clark (1969) have recently shown, in fact, that such a strategy increased the recall of twelve similar lists of words sevenfold.

Teachers are not alone in conceiving of teaching as a cause-effect, linear process. Researchers investigating teaching, for example, try to account for children's learning directly in terms of the number of moves made by the teacher, the number of questions asked, or the number of informative statements made. But the critical issue is the number of times the point gets made, no matter who is instrumental in making it. In other words, if the teacher can get the child to agree to the objective, or better still, if the teacher can find an objective that both satisfies the requirements of a general education and matches a deep current interest of the child, he may have to say or do very little, because the child will do most of the work himself. This is the conception of teaching formulated by Dewey (1902). A current example is the "Multi-Media Kit" developed by Barton and Stansfield (1970), in which novel, interesting materials completely lacking in organization are shown to a class. The class then devises a way to study the materials, developing both the objectives and a strategy for conducting their inquiry. Although the teacher is a valued consultant, he rarely provides either information or structure.

Of course, a teacher cannot simply assume that the children have the skills and motivation necessary for such a project. He must continuously monitor the children's activities, supplying critical skills and resources the children cannot supply themselves. This process requires the continuous interaction that in cybernetic terms is called feedback. A preliminary investigation of this function is the purpose of the last study in this series.

EXPERIMENT 7: THE ROLE OF FEEDBACK IN TEACHER-CHILD COMMUNICATION

In the preceding experiments, teaching has been considered more as transmission, a one-way process, than as communication, or a two-way process. The latter is derived from the former by the insertion of a feedback loop. As Wiener (1954), one of the originators of cybernetics, pointed out, "Feedback is a method of controlling a system by reinserting into it the results of its past performance [p. 48]." From this point of view, a teacher is not transmitting an objective message based on the course of study or the textbook, but rather responding to the feedback provided by the child. What the teacher does is contingent on what the child does; this relationship is just the opposite of that expressed in the Smith (1960) model, in which the teacher was the independent variable and the child the dependent variable. It is also just the opposite of the dominant contingency relationship in instructional theory, in which the teacher provides feedback or reinforcement or knowledge of results to the learner.

Feedback from the child as a controlling device in the instructional process has long been emphasized in other contexts. Augustine, with his conception of the teacher as prompter, is only one of a list that would also include Dewey (1902): "No such thing as imposition of truth from without is possible. All depends upon the activity which the mind itself undergoes in responding to what is presented from without [p. 31]." Thus the teacher is seen less as a transmitter of the culture than as an articulator of the intuitions of the child.

But research on how the activities of the teacher are determined by feedback from the child has not proceeded very far. What sensitivity to feedback from the child do teachers have? What cues do teachers pay attention to--those relevant to discipline, or boredom, or comprehension? Do teachers vary in this regard? It is a common complaint that teachers often "transmit," that is, they speak as if the lesson were independent of the learner. The responsiveness of teachers to the behavior of the children is indirectly involved in current studies of classroom interaction. Although the taxonomies of behavior used in these interaction studies have the limitation of any ad hoc taxonomy (Kierkegaard's criticism of the claim that there are three types of people--officers, serving maids, and chimney sweeps), there is some consistency in the finding that indirect teaching moves, that is, moves contingent on the child's behavior, produce both better attitudes and greater achievement than direct, noncontingent teaching moves. Campbell and Barnes (1969), after reviewing several of these studies, concluded: "The micro-elements involved in the indirect/direct ratio do affect achievement and attitude development at almost every grade level from K-9 [p. 588]." Rosenshine (1970), however, has contended that this conclusion is greatly overstated and that the relationship between indirect teaching methods and achievement remains in question.

Feedback requires a continuous monitoring by the teacher or encoder of the decoder's behavior. Specifically, feedback cues are bits of be-

havior that indicate to the encoder whether and to what extent there is a discrepancy between the message intended by the encoder and the message actually received by the decoder. Using these feedback cues, the encoder continuously tracks this discrepancy and tries to reduce it by altering his successive recodings of the message. The relationship of intended message, received message, and feedback is illustrated in Figure 7-1, in which the overlap of the two circles indicates the effectiveness of the communication, that is, the degree of correspondence between intended message and received message.

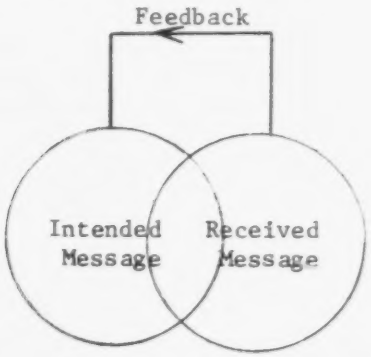

Fig. 7-1. Intended message, received message, and feedback.

Some studies have investigated more specifically the nature and effects of feedback in a communication setting. Flavell et al. (1968), reported on a series of studies of the development of the child's ability as encoder to take the role of his audience and to adjust a message to their particular requirements. They state their objective as follows:

> Its major aim was to investigate the ontogenetic development of two important forms of social-cognitive behavior: 1) the general ability and disposition to "take the role" of another person in the cognitive sense, that is, to assess his response capacities and tendencies in a given situation; and 2) the more specific ability to use this understanding of the other person's role as a tool in communicating effectively with him [p. 1].

They go on to present considerable evidence that young children's communication tends to be egocentric or nonsocial, in that young children usually do not make allowances for the requirements of the listener. Older children's communication tends to be more social--that is, better adapted to the listener--either on the basis of feedback from the listener or on the basis of previous knowledge about the listener.

Glucksberg and Krauss (1967) also presented convincing evidence of this development. For example, they found that a young child would de-

scribe the shape of a pattern to another child by drawing around it with his finger, even if the listener was separated from him by an opaque screen. They also showed that if the listener, after hearing the description, said, "I don't understand; tell me more," young children often simply repeated what they had said the first time, whereas older children and adults usually modified the message to make it more explicit. As could be expected, the older children were more effective in these communication tasks.

Hunt (1970) has reported on a series of studies that attempted to assess the social aspects of communication in various training agents, such as teachers and Peace Corps volunteers, by means of an Adaptability Index. This index is a behavioral measure of an encoder's ability to adapt his teaching to the particular background or frame of reference of the decoders. Ss were required to communicate the concept "balance of power" to a "stooge" Venezuelan immigrant who created various obstacles by misconstruing what the S said. The Ss' ability to adapt their messages to the listener's needs was positively related to some other measures of their adaptability. Moreover, there was a small but significant relationship between the adaptability of the S and the likelihood of his being successful in the training program. The studies did not examine whether this adaptability accounted for the Ss' subsequent degree of effectiveness. That is, there was no measurement of the effects on the learners of being taught by teachers with certain levels of adaptability.

The effect on the decodability of a message of giving the encoders feedback from the listeners was indicated by Leavitt and Mueller (1951). In a study of the communication of geometric forms, which served as a model for the present study, they found that restricting the feedback from the listener decreased the accuracy of the information transmitted. Furthermore, they found that free feedback appeared to permit participants to learn a shared language, whereas the zero-feedback condition was frustrating to them. The study's conclusions were weakened, however, by the fact that the sample of encoders was very small and statistical tests were applied only in the case of the extreme conditions--free feedback and zero feedback.

Still another demonstration of the importance of feedback is the finding by television producers that some actors perform better before a studio audience, apparently because they get relevant information from the audience as well as conveying information to it. If feedback from child to teacher facilitates the communication process by helping the teacher reduce the discrepancy between the intended message and the child's received message, it follows--and this is the central hypothesis of this experiment--that depriving the teacher of feedback cues will severely reduce the amount of information transmitted. In other words, our hypothesis is that the effectiveness of a communication system is a linear function of the degree of feedback incorporated in it.

The present study examined the effects of a very limited range of auditory and visual feedback cues from decoder to encoder. Auditory feed-

back was limited to the decoder's saying "yes" or "no" in reaction to what the encoder said; visual feedback was limited to the decoder's signaling comprehension or noncomprehension by "yes" or "no" gestures. These limitations were employed against the possibility that decoders might usurp the encoding function entirely; otherwise, for example, a skilled decoder might start asking questions that would reduce the encoder's role to that of merely answering "yes" or "no."

The specific hypotheses in the experiment were as follows:

1. The decoder's performance will be significantly better under conditions of visual feedback than under conditions of no feedback.
2. The decoder's performance will be significantly better under conditions of auditory feedback than under conditions of no feedback.
3. The decoder's performance will be significantly better under conditions of visual and auditory feedback combined than under either of these conditions separately.
4. The intelligence of both the encoder and the decoder will be related to the decoder's performance.
5. The time required for the communication task will be a function of the degree of feedback in the system.
6. The degree to which the encoders and decoders agree in their perceptions of both the informativeness of the messages and the enjoyability of the task will be a function of the degree of feedback in the communication system.

Subjects

The Ss for this experiment were 162 seventh graders between ten and twelve years old, half of them males and half females. The Ss were run in pairs (dyads) of the same sex. Only 43 of the dyads, or 86 Ss, succeeded in the warm-up tasks and took part in the actual experiment. These 43 dyads were assigned randomly to four treatment conditions, with ten dyads in one condition and eleven dyads in each of the remaining three.

Experimental Groups

Dyads were assigned to one of four groups in a 2 x 2 factorial design where one factor was the presence or absence of auditory feedback (A) and the other factor was the presence or absence of visual/gestural feedback (V). The combinations of these two factors yielded the four treatment groups: no feedback (NN), visual feedback only (NV), auditory feedback only (NA), both visual and auditory feedback (AV).

Materials

The information to be communicated by the encoder was how to draw a certain geometic figure consisting of a series of three connected rectangles subtended from the base of a given rectangle. The encoder viewed this figure and then attempted to describe it in such a way that the decoder could draw it without having seen it himself. Two two-rectangle patterns were used as the warm-up task; four four-rectangle patterns were used in the actual experiment. For a sample problem see Figures 6-1 and 6-2 above.

The decoders' drawings were scored for their structural similarity to the figure viewed by the encoder. Six points were given for the proper placement of each of the three subtended rectangles, as follows: two points for using the correct anchoring point on the base or previously drawn rectangle; two points for using the correct point of attachment on the present rectangle; and two points for the correct orientation of the present rectangle. Thus the maximum total score for the four problems of the communication task was 72.

Procedures

The encoder and the decoder sat facing each other at the opposite ends of a table six feet long. In the experimental conditions calling for no visual feedback, they were separated by an opaque screen four feet wide and four feet high. For the dyads in other experimental conditions, a small barrier eight inches high was placed between the encoder and the decoder, thereby permitting them to see each other's faces but not the papers lying on the table. The dialogue of each of the dyads, which was recorded on a Wollensack tape recorder, was subsequently transcribed for analysis. A stopwatch was used to determine the time taken for each interaction.

The dyads, who were sent by their classroom teacher to an empty classroom that served as a laboratory, were assigned sequentially to one of the four treatments. Upon arriving in the laboratory, the Ss were asked to take a seat. The child taking the chair at the left of E became the encoder, and the other child, the decoder. The encoder was then given the problem booklet, which contained the two warm-up problems and the four experimental problems. The decoder was given several identical answer sheets, each of which had a single rectangle at the top of the page. The encoder was shown a sample of the decoder's answer sheet and was told that he was to try to describe the patterns in such a way that the other child could draw each of them. The decoder was told that he was to listen carefully to what the speaker said and try to make a drawing like that described to him. Further instructions depended on the experimental treatment to which the dyad had been assigned. In condition NN, the decoder was told that he could say nothing. In condition NA, he was told that he could inform the speaker of his ability or inability to follow the speaker's instructions by saying either "yes" or "no." In condition NV he was told that he could say nothing but that he could shake his head to inform

the speaker of his ability or inability to follow the instructions. In condition AV he was told that he could both nod his head and say "yes" or "no" to indicate to the speaker his ability or inability to follow the instructions.

The Ss were then given the two warm-up problems. After each attempt, the encoder and decoder were permitted to compare the decoder's drawing with the model. Each dyad was given three chances to achieve success on the warm-up problems. If a pair had not succeeded at the end of this time, they were dropped from the experiment and returned to their classroom. Those who succeeded on the warm-up tasks were permitted to work through the four experimental problems. This time, however, they were not permitted to compare the drawings to the model at the end of each problem.

E tape-recorded the interactions and noted the length of time required for the four experimental tasks. Following the experiment, the encoder was asked, "Did you feel that you got your message across?" "Did you find the experience an enjoyable one?" The decoder was asked, "Did you understand what he meant?" "Did you find the experience an enjoyable one?" The total time in the experiment for each dyad was approximatley twenty minutes.

Results and Discussion

The decoders' drawings were scored by two independent judges using the scoring system outlined above. Using the Kuder-Richardson formula for computing inter-rater reliability, an r = .99 was obtained.

Three of the dyads violated the restrictions imposed on feedback, and accordingly their scores were not used in comparing the effects of the feedback conditions. The performance of the remaining 40 dyads under the different feedback conditions is shown in Figure 7-2. The significance of these differences was examined by means of a two-way analysis of variance, the results of which are given in Table 7-1. As this table shows, the communication systems that permitted auditory feedback were more effective than those that did not. Similarly, the communication systems that permitted visual (gestural) feedback were more effective than those that did not. Finally, there was a significant interaction between the audio feedback condition and the visual feedback condition. As Figure 7-2 shows, decoder performance was considerably higher when the encoder had access to both auditory and visual feedback from the decoder. These results support the study's primary hypotheses--that both auditory and visual feedback increase the level of decoder performance, and that these effects are to some extent additive.

The effects of the encoder's and decoder's intelligence were examined in three ways. First IQ, as measured by the Dominion Group Test, and school achievement, as measured by the average score on the five core subjects in the Christmas examinations, were compared for the dyads that were accepted for the experiment and those that were rejected. These two groups were not significantly different in either IQ ($t = 0.065$, $df = 38$, n.s.) or school achievement ($t = 0.36$, $df = 38$, n.s.).

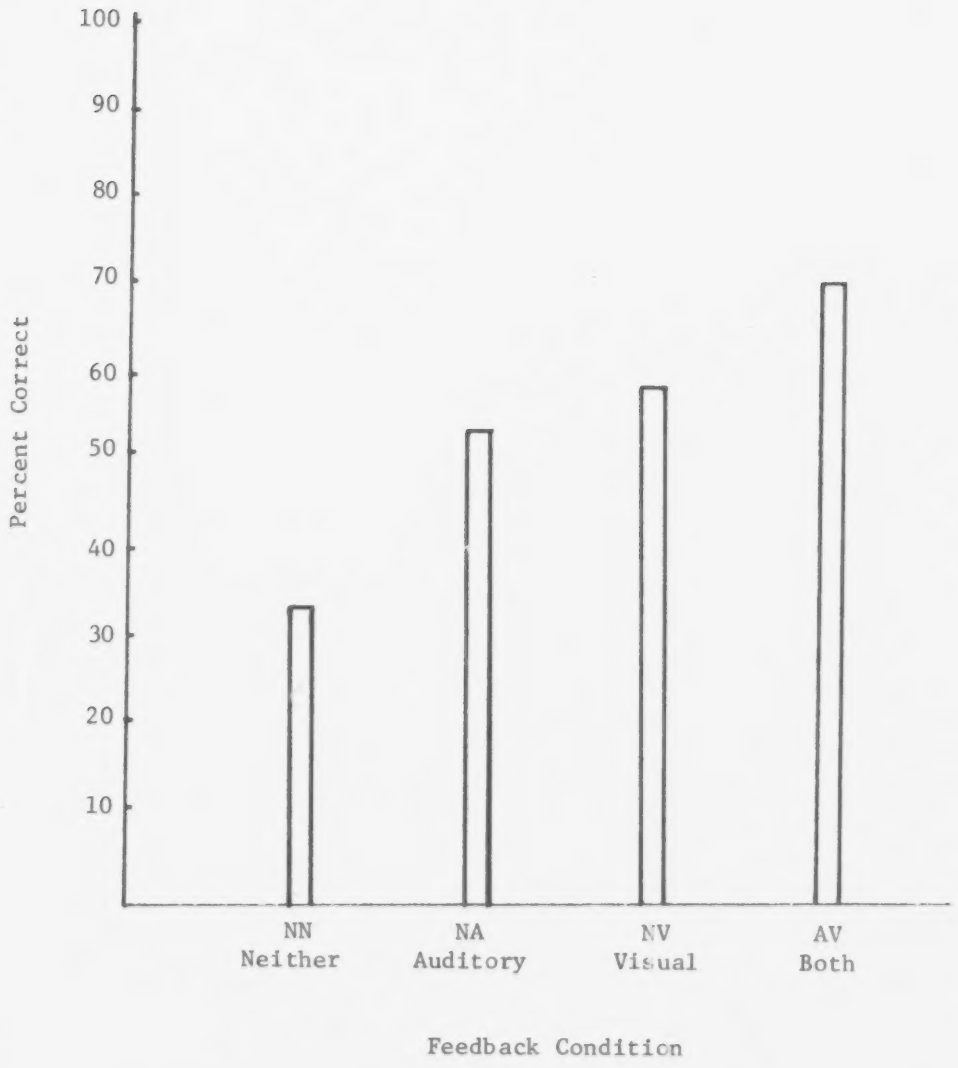

Fig. 7-2. Communication effectiveness under four conditions of feedback as measured by decoder performance.

TABLE 7-1

Two-Way Analysis of Variance Effects of Auditory and Visual Feedback

Source of Variation	SS	df	MS	F
Visual	720.00	1	720.00	4.79*
Auditory	1714.05	1	1714.05	11.41**
Visual auditory	1102.95	1	1102.95	7.34
Within sets	5404.07	36	150.11	
Total	8941.07			

*p < .05
**p < .01

Second, the average IQ of the dyads who scored above the median for their experimental group was compared to that of those who scored below the median. For all four experimental groups, the high-scoring dyads had higher IQs than the low-scoring dyads, but the differences were small, ranging from 2.0 to 6.1 IQ points. Third, an attempt was made to determine whether the decoder's degree of success was more attributable to his own IQ or to the IQ of the encoder for each of the experimental groups. This attempt involved a series of multiple correlation analyses, the results of which are presented in Table 7-2. As this table shows, whereas neither the IQ of the decoder nor that of the encoder is significantly correlated with decoder performance, presumably because of the small N, there is a slight tendency for the decoder's IQ to be more related to performance than the encoder's IQ--a pattern that was also found in Experiment 6. When all experimental conditions were combined, the multiple correlation of the encoder's and the decoder's IQ with decoder performance reaches significance. This low correlation between IQ and decoder performance may be attributable to the homogeneity of the dyads. More interesting is the possibility that high ability on the part of either member of the dyad may result in a high level of decoder performance: that is, high-ability decoders may be able to do relatively well with even fragmentary messages, and low-ability decoders may be able to do relatively well if they are given clearly formulated messages. Even poor students benefit greatly from good teaching, in other words, and good students benefit from even poor teaching. Considerable further research would be required to clarify this relationship, however.

TABLE 7-2

Multiple Correlation Coefficients for Encoder IQ, Decoder IQ, and Communication Score under Four Feedback Conditions

Coefficient	Feedback Condition				
	NN (N = 10 dyads)	NA (N = 11 dyads)	NV (N = 11 dyads)	AV (N = 11 dyads)	All conditions (N = 43 dyads)
$r_{x.z}$	0.20	0.11	0.14	-0.53	0.15
$r_{y.z}$	0.51	0.50	-0.11	0.14	0.24
$r_{xy.z}$	0.52	0.59	0.17	0.53	0.27*

Note: x = encoder IQ; y = decoder IQ; z = communication score.
*$p < .05$

The time taken for the interaction was measured and compared across the various feedback conditions. There appeared to be a tendency for conditions involving auditory feedback to take longer (NA = 17.8 minutes; AV = 14.9 minutes) than those not involving auditory feedback (NN = 12.3 minutes; NV = 7.4 minutes), but there was no obvious relationship between effectiveness and amount of time taken in the interaction.

In general, the encoders and decoders agreed in their perceptions of the enjoyability and informativeness of the exchange slightly more under the conditions permitting feedback than otherwise. The differences were not statistically significant, however, either for enjoyability ($X^2 = 6.75$, df = 3, n.s.), or for informativeness ($X^2 = 2.46$, df = 3, n.s.).

Analyzing the transcripts again proved extremely difficult. The decoders misconstrued the messages in many different ways, including some of those already indicated in the preceding experiments. Some of the tangles that our young encoders generated are illustrated by the following exerpts from transcripts:

> Encoder #17: To the rectangle there is there you must draw another rectangle except straight sort of up, touching right-hand sides. The rectangle that is there, you must draw another rectangle except it should be sort of up, like sort of vertical. It must be touching.
>
> Encoder #2: Make a base at the bottom right-hand corner of your third base on a downward angle coming down from the big base, and make an ordinary rectangle except it will be below the

rectangle below the top rectangle, and it will be kind of, there'll be a part sticking out to the left that will have to be below the rectangle coming down straight. Okay, make a rectangle on the right-hand angle with the top of the left, the left top corner has to touch the left lower corner, at the beginning of the rectangle that went across the first rectangle. Turn the page.

Encoder #7: . . . from the tip, hanging from the tip, then two lines on either side about three or four inches in. Then the third is just like the base except it's in half of number two, going straight across and down both sides, across from the bottom one.

The following exchange occurred in the warm-up session:

Decoder: Sir, I don't understand what he's saying.

Encoder: You're kidding. What more does he want, a picture?

Experimenter: Have another shot at that one and see. . . .

Encoder: Didn't he get that one either?

Experimenter: No.

Encoder: Didn't he? Okay. Now you know your base, the bottom tip, the right bottom tip, like the corner, now you take the top tip, the right-hand tip, yes, okay, I'll do it again, sorry. Forget about that, your base has a right-hand corner at the bottom. Now you bring this other rectangle, you take the right corner and it meets just in the, like both tips touch each other on the bottom of the base and you bring it down and draw your rectangle. See it's on the short top side, the long top side, and they just meet at the corner.

Decoder: I don't understand.

Experimenter: Say it again in different words.

Encoder: Different words. I don't know any other words.

Obviously these encoders found this communication task very difficult. There seem to be only a few ways of communicating this type of information effectively, and an infinite number of ways to go wrong. The major difficulties encountered by the encoders in this experiment were similar to those we described in Experiment 6: the failure to specify whether single lines or whole rectangles were the units being described; the failure to be sufficiently explicit or informative about each step in making the drawing; the failure to consider the kinds of difficulties a decoder might be having; and the inability to recode in a different way material that was puzzling the decoder. As the above transcript indicates, even when E suggested that the encoder take a different tack, using different words, one encoder replied, "I don't know any other words." Failing to recode, the encoder could only repeat his incomprehensible instruction, though frequently with more emphasis.

Conclusions

This study has clearly shown the important role played by feedback in increasing the decodability of a message. Feedback was shown to be important even when the listener had had no previous acquaintance with the information to be transmitted--a situation frequently encountered in classroom teaching. These preliminary results indicate, moreover, that within limits the degree of feedback is directly related to the effectiveness of the communication: the more feedback, the better the decoder's performance. It is reasonable to suggest that feedback is responsible for synchronizing what the teacher says with what the child is thinking, a relationship that may account for the positive findings of researchers studying classroom interactions (Amidon & Flanders, 1967; Flanders, 1970): that is, a high incidence of teacher statements that are contingent on feedback from the child would tend to generate a high "indirect" index, and it is this index that some studies have linked to improved student attitudes and achievement (Campbell & Barnes, 1969). Hence the degree to which teaching acts are contingent on feedback from the learner may be the primary factor in determining levels of achievement. If this study has not proved the existence of such an interaction, it has indicated a more simple and direct method of experimentally investigating it.

Two other hypotheses about the nature and effects of feedback on the communication of information should be mentioned. Hunt (1971) has presented evidence to the effect that adaptable encoders are more sensitive to feedback. In this study we have presented evidence that feedback increases the effectiveness of communication. These two propositions combined yield the further hypothesis, yet to be tested, that adaptable encoders are more effective in a communication situation like the one in this study. Another possibility must be taken into account, however. It may be that good teachers, being more adept at inferring the requirements of their listeners, are less dependent on feedback cues than poor teachers. In other words, there may be a complex trade-off between feedback cues directly perceived by the teacher from the present context and cues inferred from the teacher's knowledge of the types of difficulties children usually have.

One observation relevant to this trade-off is that speaking to a small group is radically different from speaking to a large group. Equally embarassing are the speaker who tries to handle a large audience as if it were a small discussion group and the speaker who lectures to a small discussion group as if it were a large audience. In the case of the small group, the speaker can be guided by immediate feedback--by gestures and expressions as well as questions from the audience--whereas in a lecture the speaker must infer or guess the interests, puzzlements, and questions of his audience. Some problems of major importance, such as that of optimal class size, cannot be solved until our knowledge of these processes is more explicit.

The second hypothesis was mentioned at the outset of this discussion: it is that different teachers monitor both different amounts and different types of feedback cues. Of all the cues that could be picked up from children in a classroom, teachers must select a very limited set. Which ones do they select--those pertaining to class discipline, to comprehension, or to mental health? And what are the differential effects on communication effectiveness of a teacher's responding to these various types of cues? Can teachers be trained to monitor a broader range of cues or to attend more selectively to those relevant to each situation? One specific type of cue has been investigated by Jecker, Maccoby, and Breitrose (1965). They found that teachers could not "read" comprehension cues from the faces of children with any better than chance levels of accuracy. They obtained some evidence, however, that training could improve these skills. What other cues teachers monitor and how accurately they interpret these cues is not known.

GENERAL SUMMARY, CONCLUSIONS, AND IMPLICATIONS

This inquiry began with the observation that learning from instruction was a radically more efficient means of obtaining information than learning by discovery. On the assumption that these two processes differ not only in their effectiveness but in their functioning, a series of experiments was conducted to determine some of the factors involved in learning from verbal instruction. The perspective adopted was that of communication theory, in which the teacher is considered an encoder who tries to generate a message that will produce a desired effect on the child, who acts as the decoder of the message.

Not surprisingly, though the encoders in our experiments were somewhat consistent in the messages they generated, they differed greatly in style, strategy, and sensitivity to their listeners. Moreover, some of these differences between encoders affected the extent to which the listeners were able to comprehend and remember the information transmitted to them. Similarly, it was found that formally alternative ways of organizing the same information--that is, recodings--were also differentially effective, and there was some evidence that some recodings were more appropriate for some children than for others. Although teachers varied considerably in their ability to recode information adequately and adjust these recodings to the requirements of their listeners, we had little success in isolating a set of ability or personality characteristics in teachers that could account for this variability.

What are we to make of these findings? Three kinds of considerations may be examined--the substantive findings themselves; the method of research developed here, that is, the usefulness of the paradigm; and the practical application of this research to improving instruction. A summary model of the process investigated in the experiments is shown in Figure 8-1. According to this model, both the teacher and the child have goals, expectations, and intentions. Incoming information either from the perceptual world or from the actions or speech of the other is compared with these expectations and intentions; in the event of a mismatch, some readjustment occurs in the form of a recoding that in turn yields some performatory act, such as doing, making, or saying something. At this point the activities of the teacher and student frequently diverge, owing to the indifferent information resources. The teacher formulates a message conveying the information to be learned by the child; the child formulates some paraphrase or carries out some other act according to his perception of this information. Both the teacher and the child monitor this act to see whether it corresponds to their intentions. If it does, they go on to a new topic; if not, the system recycles, with a new recoding of the information in a hopefully more appropriate form that takes account of the preceding performatory acts.

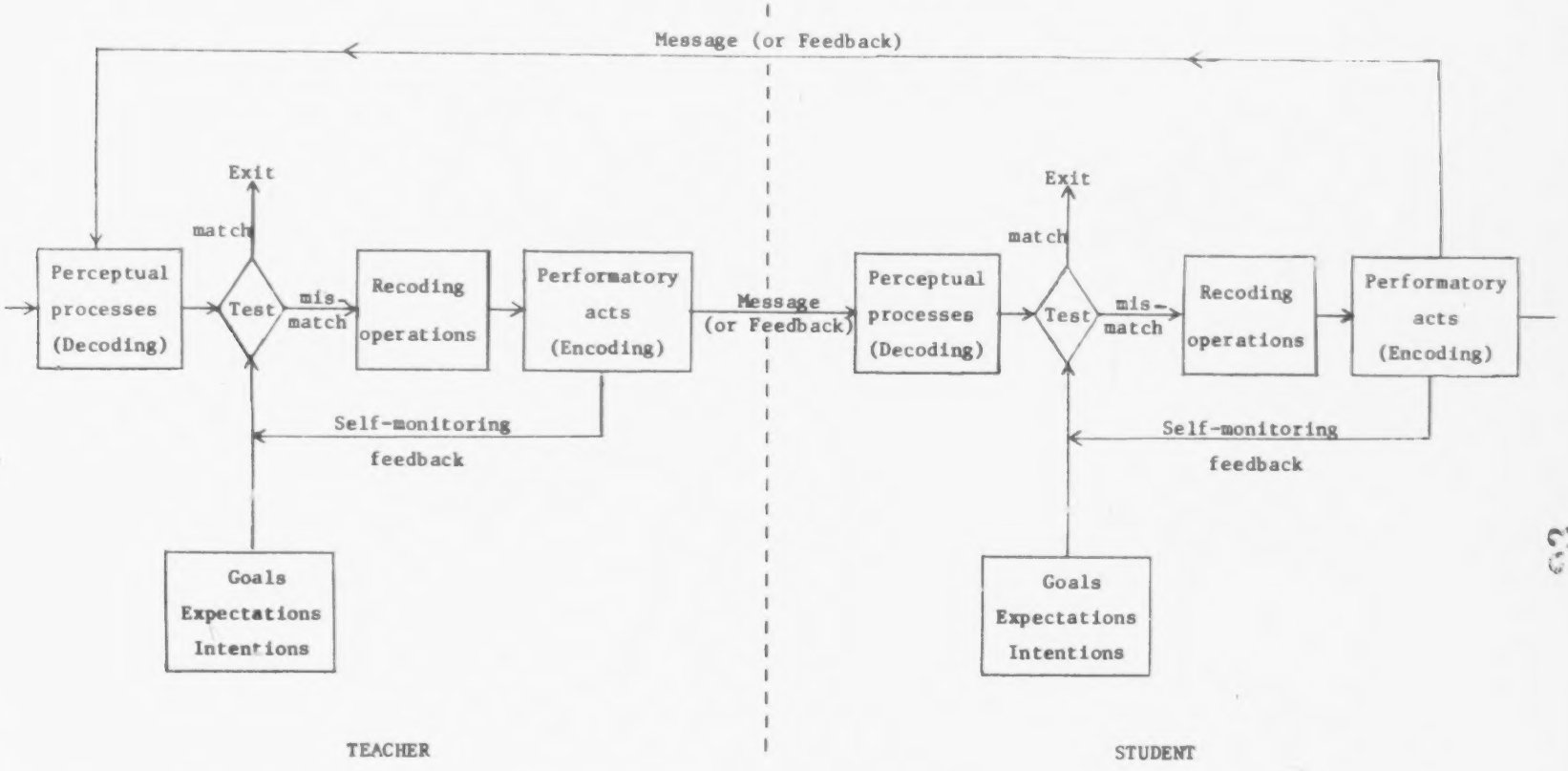

Fig. 8-1. A schematic model of the teacher-student communication system.

Thus teacher-student interaction, according to this model, is symmetrical in several respects. At different points in time and with different forms of instruction, either child or teacher may act as the encoder, feedback from the child being the counterpart of the teacher's message. Moreover, either child or teacher may monitor his own performance and that of the other, and judge the appropriateness of these results to his own intentions. As noted earlier, however, the teacher can rely on feedback directly from the child's performance only if the class is small; otherwise, the teacher, as lecturer, must rely on his own anticipation of probable feedback and on monitoring his own encoding acts.

We may now turn to the substantive findings of these experiments. In investigating the effects of various recodings, the experimenters designed lessons or messages that contained the same information but organized it in different ways. These different recodings were found to produce different degrees of learning and recall in children. Although certain means of recoding may be more effective than others for certain types of information, it appeared in general that the most highly structured recodings--that is, those in which each piece of information was tied into the other pieces of information--were the most effective. Of course, structuring information is the whole function of theories; its importance was emphasized by Bruner (1960) in his report on teaching the structure of a discipline. As he pointed out, a subject should be taught in a way that will produce "the most fundamental understanding that can be achieved of the underlying principles that give structure to that subject [p. 31]." This structure, he went on to say, gives coherence to the discipline, thereby providing for a high degree of recall and transfer.

All of the recodings examined in the present studies were communicated by language. It may be, however, that some ideas are best communicated through other media, such as drawings or film. This possibility was not examined specifically in these studies. It was obvious, however, that both teachers and children had great difficulty in rendering the geometric drawings verbally and in extracting the geometric information from the verbal presentations. But the high rate of failure on this task reflects less on the capacities of the teachers and children than on the use of verbal means for transmitting geometric information. Had pictures or blackboard drawings been used, the performance would undoubtedly have improved markedly. In school, children who fail to extract information about the structure of molecular lattices from a brief verbal description may often be considered dull, when the fault is more likely that of the textbook writer and/or the teacher, who simply failed to use an appropriate medium, such as film, for presenting spatial-visual information. In the same way that various recodings in language may account for different degrees of learning and recall, so recodings into other media may also have a marked effect.

In addition to considering the type of information to be recoded, one must consider the type of child expected to learn the information.

Some types of recodings appear to be more appropriate for some kinds of learners than for others. It is this possibility that inspired Bruner (1960) to write: "We begin with the hypothesis that any subject can be taught effectively in some intellectually honest form to any child at any stage of development [p. 33]." That is, any piece of information can be recoded in several different ways; these different recodings may be differentially appropriate to different children at different stages of development. The present studies support to some extent the idea that young, preoperational children need to have information organized in a different way than older children, but that important problem requires considerably more research. The relationship between the type of recoding or form of instruction and the requirements of the listener is also the basis of Cronbach and Snow's (1969) concept of aptitude-treatment interaction (ATI). However, the substantial effects attributable to different types of recodings, compared to the small effects attributable to ability differences among the listeners, lead us to reconsider this interaction. Perhaps ATI is an aspect of increased contingency between what the child is thinking and what the teacher is doing. This line of interpretation is supported primarily by the substantial effects on the decodability of a message of the teacher-encoder's continuous monitoring of feedback from the learner. This feedback was interpreted as a means of synchronizing the activities of the teacher and learner--a means of making the teacher's acts contingent on the child's behavior. Matching aptitude with treatment may be an attempt to accomplish from a distance the very thing that is accomplished directly in effective communication; that is, it may be an attempt to provide for the recoding of the message in different ways until it suits the informational requirements of the listener, who then signals that match by feedback. Certainly the synchronizing effect of feedback was one of the most critical factors in these studies of instruction.

Is there such a thing as encoding ability, and is it related to teaching ability? Some of our encoders were found to be consistently superior to others both in teaching a message initially and in reteaching a message. It would be of great interest to know whether this consistency was maintained over a wide range of communication tasks and whether skill in performing these tasks was related to skill in the classroom. Further research efforts should be devoted to developing a battery of brief lessons that could be taught by student-teachers and subsequently evaluated. Such a set of tasks might be useful both in selecting teacher-trainees and in training teachers. This suggested procedure has some things in common with the current use of "microteaching" as a training device (Allen & Ryan, 1969). It remains to be seen whether the ability to communicate can be developed by training procedures.

The components of a skilled teaching act have only been touched on in this series of studies. Those examined, however, include the accurate assessment of the listener's requirements in terms of his expectations, his frame of reference, and his stage of development; the use of optimal recodings or structuring of the material; and sensitivity to feedback from the listener. Obviously this list is far from being a complete catalog of teaching skills.

Next we must assess the usefulness of the communication-games paradigm. Apart from their general relevance to the communicative aspects of teaching, the communication games permitted us to examine some of the specific communication processes, including encoding, decoding, encoder-decoder interactions, and feedback. While it cannot be argued that these processes are identical in two-person communication games and in classroom interactions, there appears to be enough similarity to justify using this paradigm to examine some of the fundamental aspects of teaching. Once the parameters of the teaching-communicating process have been isolated in the relatively simple communication games, it may be reasonable to examine them in the larger, more complex environment of the classroom. Without the benefit of preliminary work in a simpler context, the researcher in the classroom has difficulty deciding what to observe and measure. Moreover, it does seem to be the case that many of the factors detected in long-term, costly, and elaborate classroom research have been isolated, manipulated, and examined in the short-term, relatively simple studies conducted here.

As for the practical application of the findings reported here, we are less optimistic. It is not at all clear that teachers will improve in their skills as a result of knowing about these experiments. For example, the types of recodings that were theoretically derived and examined here turned out to be very similar to the forms of instruction the teachers devised when they were given the encoding task. Apparently they already knew from experience or by intuition the factors that affect list learning and the positive effects of organization in recall. These devices were certainly known in the practical arts of memory improvement. Furthermore, clever teachers, given their freedom, could probably think up more effective recodings than we were able to derive theoretically. The disadvantage of this procedure is that the researcher cannot precisely describe, at least in psychological terms, what the teacher has done. To some extent, then, the researcher and the teacher are at odds. In another context, Olson (1970a) suggested that the effectiveness of a lesson varied inversely with the precision with which it could be described. It is also a fact that the experimental psychologist is not equipped to devise the most general recodings of information: recoding knowledge into comprehensive theoretical systems is the job of the scholars in the discipline involved. The teachers in our experiments were also aware, at least to some extent, of the need to adapt their messages to their listeners' requirements. Most encoders felt severely deprived in our experiments if they were not given access to feedback. They had so much difficulty encoding without a listener that we were obligated to use a live decoder in most of the later experiments.

In spite of the fact that these studies contribute more to the theoretical analysis of teaching than to any immediate development in classroom methods, they may have some practical value in focusing the teacher's attention on some of the critical communication skills. For example, teachers would probably benefit from being made more aware of their dependence on feedback, of the type of cues they should monitor, and of the potential advantages of monitoring these cues. We suggested earlier that some

teachers may respond primarily to cues pertaining to classroom order and discipline, whereas others respond to cues relevant to the mental health of the children and only occasionally to those relevant to the comprehension of information; perhaps there is some optimal manner of distributing a teacher's attention among these types of cues. Perhaps the issue is more one of seeing which cues are most important at any one moment or in any one context.

One observation that may have some practical value is that teachers tended to take the primary responsibility for the child's learning, organizing and repeating the information for the child, whose role was thus a relatively passive one. Apparently the teachers assumed that teaching-learning was a direct cause-effect relationship. It may be important, however, for teachers to pass some of the responsibility for learning back to the child, letting him organize and rehearse the information as his capacities permit. The teacher's role is then to help him direct his attention to ideas that are both interesting and significant, and to assist him when his skills are inadequate to the learning task.

As for instruction in general, language seems a limited means, confined as it is to specifying states that are already in the system. Pauling's description conveyed the information about his discovery only to those who already knew a great deal about the field of research. We are left with only a limited understanding of instruction in general, and of verbal instruction in particular.

REFERENCES

Ackerman, W. I. Teacher competence and pupil change. Harvard Educational Review, 1954, 24, 213-89.

Allen, D., & Ryan, K. Microteaching. Reading, Mass.: Addison-Wesley, 1969.

Amidon, E. J., & Flanders, N. A. The role of the teacher in the classroom. Minneapolis: Association for Productive Teaching, 1967.

Aquinas, St. Thomas. De Veritate (Concerning truth). In K. Price (Ed.), Education and philosophical thought. Boston: Allyn and Bacon, 1967.

Augustine, St. De Magistro (The teacher). In K. Price (Ed.), Education and philosophical thought. Boston: Allyn and Bacon, 1967.

Ausubel, D. P. The psychology of meaningful verbal learning. New York: Grune and Stratton, 1963.

Ausubel, D. P. Educational psychology: A cognitive view. Toronto: Holt, Rinehart and Winston, 1968.

Barton, A., & Stansfield, D. The 1930's multi-media kit. Toronto: Ontario Institute for Studies in Education, 1970.

Berlo, D. K. The process of communication: An introduction to theory and practice. Toronto: Holt, Rinehart and Winston, 1960.

Bloom, B. Learning for mastery. Los Angeles, Calif.: UCLA Center for the Study of Evaluation of Instructional Programs. 1968, 1, (2).

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. Journal of General Psychology, 1953, 49, 229-40.

Bower, G. H. Educational applications of mnemonic devices. Mimeo., Stanford University, Department of Psychology, 1970.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 323-43.

Bower, G. H., & Clark, M. C. Narrative stories as mediators for serial learning. Psychonomic Science, 1969, 14, 181-82.

Bruce, D., & Cofer, C. N. An examination of recognition and free recall as measures of acquisition and long term retention. Journal of Experimental Psychology, 1967, 75, 283-89.

Bruner, J. S. The process of education. Cambridge: Harvard University Press, 1960.

Campbell, J. R., & Barnes, C. W. Interaction analysis--a breakthrough? Phi Delta Kappan, 1969, 50, 587-90.

Clark, H. H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 74, 387-404.

Clark, H. H., & Chase, W. G. On mental comparisons of sentences and pictures. Unpublished manuscript, Stanford University, 1970.

Clifford, T. Simplified computational programming for analysis of variance designs with correlated observations. Psychological Bulletin, 1968, 69, 439-40.

Cronbach, L. J., & Snow, R. E. Final report: Individual differences in learning ability as a function of instructional variables. Stanford, Calif.: Stanford University, School of Education, 1969.

Dewey, J. The child and the curriculum. Chicago: University of Chicago Press, 1902.

Dewey, J. The school and society. Chicago: University of Chicago Press, 1915.

Donaldson, M., & Wales, R. On the acquisition of some relational terms. In J. R. Hayes (Ed.), Cognition and the development of language. New York: Wiley, 1970.

Flanders, N. A. Analyzing classroom interaction. New York: Addison-Wesley, 1970.

Flavell, J. H. The developmental psychology of Jean Piaget. Princeton, N.J.: Van Nostrand, 1963.

Flavell, J. H., Beach, D. R., & Chinsky, J. M. Spontaneous verbal rehearsal in a memory task as a function of age. Child Development, 1966, 37, 283-99.

Flavell, J. H., Botkin, P. T., Fry, C. L., Wright, J. W., & Jarvis, P. E. The development of role-taking and communication skills in children. New York: Wiley, 1968.

Fortune, J. C. A study of the generality of presenting behaviors in teaching preschool children. U.S. Office of Education Project No. 6-8468. Memphis, Tenn.: Memphis State University, 1967.

Frase, L. T. Structural analysis of the knowledge that results from thinking about text. Journal of Educational Psychology Monograph, 1969, 60, (6), Part 2, 1-16.

Gage, N. L. Paradigms for research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 94-141.

Gagné, R. M., & Rohwer, W. D. Instructional psychology. Annual Review of Psychology, 1969, 20, 381-418.

Glucksberg, S., & Krauss, R. M. What do people say after they have learned to talk?: Studies of the development of referential communication. Merrill-Palmer Quarterly, 1967, 13, 309-16.

Gough, P. B. Grammatical transformations and the speed of understanding. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 107-11.

Harvey, O. J., Hunt, D. E., & Schroder, H. M. Conceptual systems and personality organization. New York: Wiley, 1961.

Hunt, D. E. A conceptual systems change model and its application to education. In O. J. Harvey (Ed.), Experiencing, structure, and adaptability. New York: Springer, 1966. Pp. 277-302.

Hunt, D. E. A conceptual level matching model for coordinating learner characteristics with educational approaches. Interchange, 1970, (3), 68-82.

Hunt, D. E. Matching models in education: The coordination of teaching methods with student characteristics. Ontario Institute for Studies in Education, 1971.

Hunt, D. E., & Joyce, B. R. Teacher trainee personality and initial teaching style. American Educational Research Journal, 1967, 4, 253-59.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning. Psychological Review, 1968, 75, 550-60.

Inhelder, B., & Piaget, J. The early growth of logic in the child. London: Routledge, 1964.

Jecker, J. D., Maccoby, N., & Breitrose, H. S. Improving accuracy in interpreting nonverbal uses of comprehension. Psychology in the Schools, 1965, 2, 239-44.

Johnson-Laird, P. N. On understanding logically complex sentences. Quarterly Journal of Experimental Psychology, 1969, 21, 1-13.

Johnston, N. M. E. An experimental investigation into the role of language in discrimination, memory and transfer in preschool children. Unpublished doctoral dissertation, University of Toronto, 1969.

Joyce, B. R., & Harootunian, B. The structure of teaching. Chicago: Science Research Associates, 1967.

Katz, E., & Lazarsfeld, P. F. Personal influence. New York: Free Press, 1955.

Leavitt, H. J., & Mueller, A. Some effects of feedback on communication. Human Relations, 1951, 4, 401-10.

Maccoby, E. E. What copying requires. In D. R. Olson & S. Pagliuso (Eds.), From perceiving to performing: An aspect of cognitive growth. Special Issue: Ontario Journal of Educational Research, 1968, 10, 3), 155-237.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation: Advances in research and theory. New York: Academic Press, 1966.

Mandler, G. Verbal learning. In T. M. Newcomb (Ed.), New directions in psychology. Vol. 3. New York: Holt, Rinehart and Winston, 1967.

McLachlan, J. F. Individual differences and teaching methods in student interpretation of modern art. Unpublished master's thesis, Ontario Institute for Studies in Education, 1969.

Miller, G. A. The magic number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Miller, G. A. Psychology, the science of mental life. New York: Harper and Row, 1962.

Morton, J. Interaction of information on word recognition. Psychological Review, 1969, 76, 165-78.

Murdoch, I. Sartre: Romantic rationalist. New Haven, Conn.: Yale University Press, 1953.

Murdock, B. B. Recent developments in short-term memory. British Journal of Psychology, 1967, 58, 421-33.

Nuthall, G. An experimental comparison of alternative strategies for teaching concepts. American Educational Research Journal, 1968, 5, 561-84.

Olson, D. R. Cognitive development: The child's acquisition of diagonality. New York: Academic Press, 1970. (a)

Olson, D. R. Language acquisition and cognitive development. In C. Haywood (Ed.), Socio-cultural aspects of mental retardation. New York: Appleton-Century-Crofts, 1970. (b)

Olson, D. R. Language and thought: Aspects of a theory of semantics. Psychological Review, 1970, 77, 257-73. (c)

Olson, D. R., Hildyard, A., & Crosbie, D. Does the structure of verbal memory change with the onset of concrete operations? (In preparation.)

Paivio, A. Mental imagery in associative learning and memory. Psychological Review, 1969, 76, (3), 241-63.

Piaget, J., Inhelder, B., & Szeminska, A. The child's conception of geometry. London: Routledge & Kegan Paul, 1960.

Premack, D., & Schwartz, A. Preparations for discussing behaviorism with chimpanzees. In F. Smith & G. A. Miller (Eds.), The genesis of language: A psycholinguistic approach. Cambridge: M.I.T. Press, 1966. Pp. 295-336.

Quine, W. V. From a logical point of view. Cambridge: Harvard University Press, 1953.

Rohwer, W. D. Jr. Images and pictures in children's learning: Research results and instructional implications. Psychological Bulletin, 1970, 73, 393-403.

Rosenshine, B. The consistency of teacher effects. Unpublished manuscript, School of Education, Temple University, 1969.

Rosenshine, B. Interaction analysis: A tardy comment. Phi Delta Kappan, April 1970, 445-46.

Rothkopf, E. Z. Incidental memory for location of information in text. Journal of Verbal Learning and Verbal Behavior, in press.

Shannon, C., & Weaver, W. A mathematical theory of communication. Urbana: University of Illinois Press, 1949.

Simon, A., & Boyer, G. (Eds.) Mirrors for behavior: An anthology of observation instruments. Vols. 1-14. Philadelphia: Research for Better Schools, 1966-70.

Slobin, D. I. Grammatical transformations in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-27.

Smith, B. O. A concept of teaching. Teachers' College Record, 1960, 61, 229-41.

Smith, B. O., Meux, M., Coombs, J., Nuthall, G., & Precians, R. A study of the strategies of teaching. Urbana: University of Illinois, Bureau of Educational Research, 1967.

Snow, R. E. Research on media and aptitudes. In G. Salomon & R. E. Snow (Eds.), Commentaries on research in instructional media: An examination of conceptual schemes. Bulletin of the School of Education (Indiana University), 1970, 46, (5), 63-89.

Taylor, W. L. Cloze procedure: A new tool for measuring readability. Journalism Quarterly, 1953, 30, 415-33.

Tomlinson, P. D. Differential effectiveness of three teaching strategies for students in varying conceptual levels. Master's thesis, Ontario Institute for Studies in Education, 1969.

Travers, R. M. Essentials of learning. New York: Macmillan, 1963.

Tulving, E. Subjective organization in free recall of "unrelated" words. Psychological Review, 1962, 69, 344-56.

Tulving, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), Verbal behavior and general behavior theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Wanner, H. E. On remembering, forgetting, and understanding sentences: A study of the deep structure hypothesis. Unpublished doctoral dissertation, Harvard University, 1968.

Wiener, N. The human use of human beings. New York: Anchor Books, 1954.

Wittgenstein, L. Philosophical investigations. Oxford: Blackwell, 1958.

Wright, P. Transformations and the understanding of sentences. Language and Speech, 1969, 12, 156-66.